ISOLATION AND CHARACTERIZATION OF THE PUTATIVE MIH GENE SEQUENCE FROM THE LOBSTER *JASUS EDWARDSII*

A thesis
submitted in fulfilment
of the requirements for the degree
of
Doctor of Philosophy
in the
Department of Zoology

by

Jenny Gaik Imm Khoo

University of Canterbury
1996

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The X-organ sinus gland complex in the crustacean eyestalk is an important source for several neuropeptides, among which is the moult-inhibiting hormone (MIH).

The aim of this thesis was to isolate and characterize the gene for putative MIH from the lobster *Jasus edwardsii*. Previous studies have demonstrated an immunological and functional relationship between vasopressins and the moult-inhibiting hormone of crustaceans. Using primers based on the rat vasopressin gene, I was able to amplify by polymerase chain reaction, a 947 bp sequence from DNA of the lobster *Jasus edwardsii*. Northern blot analysis showed that the 947 PCR sequence could detect mRNAs extracted from the epithelia, eyestalk, heart, hepatopancreas and muscle, but one mRNA was predominantly expressed in the eyestalk. Sequence analysis of the 947 PCR product revealed an intron/exon splice junction, a protein coding region, and a stretch of repetitive sequences which translated into a metallothionein-like protein. By *in situ* hybridization, the 947 PCR sequence detected mRNAs synthesized in the neurosecretory regions of the eyestalk, and the expression of these genes appeared to be related to the moult cycle. Using this PCR product as a probe, three different cDNA clones, peJK1, peJK2 and peJK3, were isolated from a lobster eyestalk cDNA library. *In situ* hybridization studies and northern blot analysis showed that peJK2 and peJK3 were expressed predominantly in the eyestalk and to a lesser degree in the epithelia, whereas peJK1 was expressed in the epithelia, eyestalk, heart, hepatopancreas, and muscle tissue of the eyestalk. Sequence analysis of the cDNA clones showed 96.6% homology between peJK2 and peJK3, and between 44-51% sequence identity with published putative MIH cDNA sequences from three other species, *Carcinus maenas*, *Callinectes sapidus*, and *Penaeus vannamei*. Analysis of the deduced amino acid sequences of peJK2 and peJK3 revealed the presence of a signal peptide which is characteristic of secretory proteins. Sequence alignment of the deduced amino acid sequences with other

eyestalk neuropeptides of the CHH/MIH/VIH family suggested that the cDNA clones isolated from *J. edwardsii* contained novel sequences which were different from this group of peptides. Based on these results, peJK2 and peJK3 do not appear to code for the putative MIH neuropeptide as found in other crustaceans, but may be coding for a MIH-like peptide.

CHAPTER I

Introduction

The crustacean eyestalk is an important site of production for several endocrine factors, which are implicated in almost every aspect of crustacean physiology including pigment concentration and dispersion, inhibition of moulting, gonad development, and regulation of glucose levels in the haemolymph (Quackenbush, 1986; Webster and Keller, 1988).

To increase in size, crustaceans have to replace their confining exoskeleton with a larger one and subsequently grow to fill it. The new exoskeleton is formed underneath the outer exoskeleton and is secreted by a single layered epidermis. This process of periodically shedding the exoskeleton is termed moulting and is under the immediate control of the moulting hormone, 20-hydroxyecdysone (20-HE). The precursor to 20-HE is ecdysone which is synthesized and secreted by the moulting gland (Y-organs) (Fig. 1.1). The most widely accepted hypothesis concerning the control of moulting in crustaceans proposes that an increase in ecdysteroid synthesis and consequently haemolymph titre is necessary to initiate and sustain proecdysis. There appear to be also other factors, both inhibitory and stimulatory, that regulate moulting (for reviews see Aiken, 1980; Chang, 1985; Skinner, 1985; Quackenbush, 1986; Chang *et al.*, 1993; Lachaise *et al.*, 1993; Chang, 1995).

The earliest observations implicating the eyestalk neurosecretory system in the control of moulting were made by Zeleny (1905, as cited in Lachaise *et al.*, 1993), demonstrating that eyestalk ablation frequently resulted in a dramatic acceleration of moulting. The role of the X-organ and of the sinus gland in moult control was established by Passano (1953), as cited in Huberman and Aguilar (1989), who also showed that the neurosecretory factor produced was inhibitory. Since then, it has been termed the moult-inhibiting hormone (MIH).

1.1 Crustacean moult-inhibiting hormone (MIH)

In decapod crustaceans, the synthesis and/or secretion of ecdysone by the Y-organs appears to be inhibited by the neuropeptide MIH (Fig. 1.1) (Soumoff and O'Connor, 1982; Mattson and Spaziani, 1985b). MIH also acts by inhibiting 20-HE action at the level of the epidermal tissue, as shown by apparent localized responses of epidermal tissue to temporary concentration gradients of MIH and 20-HE established in the same animal (Freeman and Bartell, 1976). Freeman and Costlow (1979) also reported that aqueous extracts of shrimp eyestalks (i.e. MIH) interfered with ecdysterone-induced apolysis in barnacle mantle explants; the interference could be overcome by administering higher levels of ecdysterone. The MIH originates in the eyestalk neurosecretory cells ("X-organs") in the proximal ganglion of each eyestalk and is released from the sinus gland into the haemolymph (reviewed in Quackenbush, 1986).

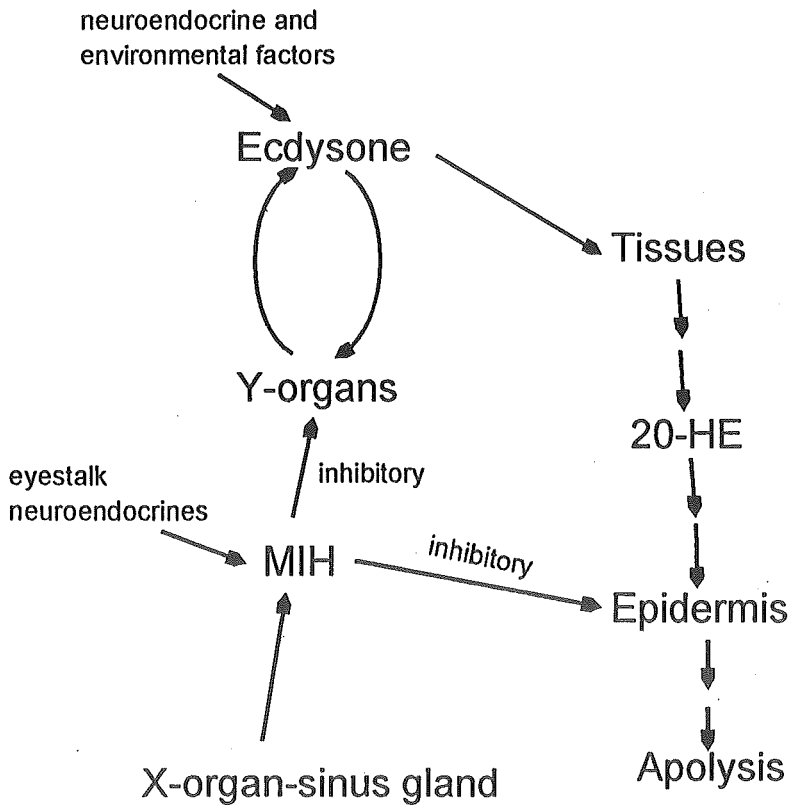


Figure 1.1. The role of moult-inhibiting hormone (MIH) in the control of moulting in crustaceans. 20-HE = 20-hydroxyecdysone (modified from Quackenbush, 1986).

1.2 Putative crustacean MIH

Although a lot of studies have been done on characterizing MIH activity, i.e. the nature and mode of function, MIH is still designated as putative as its precise biological role *in vivo* has not been elucidated. Two laboratories have reported that crustacean eyestalks contain non-peptide compounds with MIH-like activity. These have been isolated and identified as 3-hydroxy-L-kynurenine and its metabolite, xanthurenic acid (Soyez and Kleinholz, 1977; Naya *et al.*, 1988; Naya *et al.*, 1989). However, there is more evidence to suggest that MIH is proteinaceous. It is inactivated by trypsin (Rangarao, 1965; Webster, 1986), and it is a heat stable peptidic substance (Freeman and Bartell, 1976; Webster 1986).

Putative MIH neuropeptides, identified by virtue of their ability to repress ecdysteriodogenesis by Y-organs cultured *in vitro*, have been isolated from the sinus glands of several crustaceans (grass shrimp, spiny lobster, shore crab, lobster). However, the molecular weights of these polypeptides vary from species to species (Table 1.1). The general method of isolation has been gel filtration (Sephadex G-25, G-50) or high performance liquid chromatography (HPLC). The molecular weights of putative MIH peptides have been determined by elution time through the column and by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Gel filtration measurements of semipurified MIH suggest molecular weights of 1000 to 5000 (relative molecular mass (M_r)) in the grass shrimp, *Palaemonetes pugio* (Freeman and Bartell, 1976), 3400 in the spiny lobster, *Panulirus argus* (Quackenbush and Herrnkind, 1983), and between 6000 to 14000 in the shore crab, *Carcinus maenas* (Webster, 1986).

Huberman and Aguilar (1989) isolated a neuropeptide with moult inhibiting activity from the Mexican crayfish, *Procambarus bouvieri*. Based on migration on SDS-PAGE, amino acid analysis and end group analysis, they obtained a molecular weight of 6200 to 6400; i.e., 53 to 55 amino acid residues with isoleucine at the C-terminus. Chang *et al.* (1990) published the first amino acid sequence of a crustacean moult-inhibiting hormone (Hoa-MIH, 71 residues) from the lobster *Homarus americanus*. Putative moult-inhibiting hormone (Cam-MIH) has also been isolated

from the shore crab *C. maenas* (Webster, 1991). This 78 amino acid residue peptide has low homology (25%) to Hoa-MIH.

Table 1.1. Neuropeptides with moult-inhibiting hormone activity, isolated from the eyestalk of several crustacean species.

Animal	Molecular weight (M_r)	Amino acid residues	Source
Grass shrimp *P. pugio*	1000-5000	-	Freeman and Bartell (1976)
Spiny lobster *P. argus*	3485	-	Quackenbush and Herrnkind (1983)
Shore crab *C. maenas*	7200	61 a.a.	Webster and Keller (1986)
	9181	78 a.a. sequenced	Webster (1991)
Mexican crayfish *P. bouvieri*	6200-6400	53-55 a.a	Huberman and Aguilar (1989)
Lobster *H. americanus*	8700 ± 1000	60-70 a.a	Chang *et al.* (1987)
		71 a.a sequenced	Chang *et al.* (1990)

In the past few years, molecular cloning techniques have been used to isolate cDNA clones encoding the putative moult-inhibiting hormone from several crustaceans. Using PCR (polymerase chain reaction)-based screening and colony hybridization, Klein *et al.* (1993b) isolated and sequenced a cDNA encoding the complete precursor of the putative moult-inhibiting hormone of the shore crab *C. maenas*. The precursor consists of a putative 35 amino acid signal peptide and the 78 amino acid mature MIH. The sequence of the hormone, as deduced from the cDNA, confirmed the sequence obtained previously by Edman degradation of the mature polypeptide (Webster, 1991).

Probes were also generated by PCR, based on published amino acid sequences of MIH from both the lobster *H. americanus* and the shore crab, to screen cDNA libraries of the white shrimp, *Penaeus vannamei* (Sun, 1994), and the blue crab,

Callinectes sapidus (Lee *et al.*, 1995), respectively. The deduced amino acid sequence (78 residues) of *C. sapidus* MIH is 78% homologous to that of *C. maenas*. On the other hand, the deduced amino acid sequence of *P. vannamei* MIH-like neuropeptide (72 residues) shares 49% and 29% amino acid sequence identity to the MIH from *H. americanus* and *C. maenas*, respectively.

Despite the numerous studies focused on characterizing MIH, the information on the structure of the MIH polypeptide and gene is limited and sometimes contradictory. Some of the polypeptides isolated have multiple biological activities; Hoa-MIH has both MIH and crustacean hyperglycemic hormone (CHH) activity (Chang *et al.*, 1990). It has approximately 61% sequence homology to the crustacean hyperglycemic hormone (Cam-CHH, 72 a.a.) isolated from the shore crab, *C. maenas* (Kegel *et al.*, 1989; Weidemann *et al.*, 1989). The *P. vannamei* MIH-like neuropeptide appears to be more closely related to CHH than to MIH. Amino acid sequence data places MIH in a novel family of large neuropeptides from the X-organ-sinus gland complex, which includes CHH, putative MIH, and vitellogenesis-inhibiting hormone (VIH) (reviewed in Keller, 1992). In addition, there has been immunological evidence presented which suggests that MIH is related to the vasopressins (Mattson and Spaziani, 1985a).

1.3 Crustacean MIH is structurally related to the vasopressins

Using antisera raised against vertebrate neuropeptides, substance P- and enkephalin-like peptides have been localized in the eyestalk of the lobster *Panulirus interruptus* (Mancillas *et al.*, 1981). In 1982, Van Herp and Bellon-Humbert, using immunocytochemical techniques, localized vasopressin-like and neurophysin-like peptides in the eyestalk of the shrimp *Palaemon serratus*. Neurophysin is the carrier protein for the vertebrate hormone, vasopressin (Schmale *et al.*, 1983). Subsequently, researchers began to look for possible relationships between vertebrate neuropeptides and crustacean MIH.

Mattson and Spaziani (1985a) tested the effects of several vertebrate neuropeptides for possible MIH-like biological activity *in vitro* on cultured Y-organs

from the crab *Cancer antennarius*. They found that peptides of the vasopressin-oxytocin family mimicked MIH action by inhibiting Y-organ ecdysteroid production with relative potencies: lysine vasopressin (LVP) > arginine vasopressin (AVP) > vasotocin (VT) >>> oxytocin (OT). When they compared the effects of LVP and MIH activity from sinus gland extracts of the crab *C. antennarius*, on cultured Y-organs, the following similarities were observed. Maximal doses of MIH and LVP inhibited ecdysteroidogenesis with similar time courses; the inhibitory effect of LVP on Y-organs was reversible as demonstrated with MIH (Mattson and Spaziani, 1985b); and, like MIH, LVP induced significant increases in Y-organ cAMP content to 200-250% of control levels (Mattson and Spaziani, 1985c).

The results of a comparative study of the immunological properties of MIH (from sinus gland extract) and vasopressin related peptides suggest that MIH is most closely related to LVP (Mattson and Spaziani, 1985a). LVP and MIH had similar affinity characteristics with antibody to AVP, in contrast to vasotocin and oxytocin; the competitive binding curves for vasotocin and oxytocin with antibody to AVP were distinct from LVP and MIH.

In agreement with these observations, identical doses of AVP antiserum were required to attenuate the MIH- and LVP- induced suppression of ecdysteroidogenesis. Both MIH- and LVP-induced inhibition of ecdysteroidogenesis was significant at AVP antiserum dilutions of less than 10^{-4} M, whereas the AVP-induced inhibition was significant at AVP antiserum dilutions of less than 10^{-6} M. Cross reactivities of MIH, LVP, VT and OT with the AVP antiserum were 41%, 37%, 2% and <0.1%, respectively. Cross reactivities were calculated by dividing the ED50 (AVP antiserum required to attenuate the peptide inhibitory effect by 50%) for AVP by that of the given peptide tested. Ecdysteroid production was measured by the radioimmunoassay technique (Mattson and Spaziani, 1985b).

1.4 Vasopressin-oxytocin homologues

With the advent of immunological and molecular biology techniques, several polypeptides have been identified and isolated from invertebrate sources that are

homologous to vertebrate polypeptides (reviewed in De Loof and Schoofs, 1990). Amongst these is the vasopressin-oxytocin family.

The vasopressin-oxytocin hormone superfamily consists of vasopressin, oxytocin, and related peptides. In vertebrates, the hormones vasopressin and oxytocin are involved in the processes of water resorption and uterine contraction, respectively (Richter, 1987). Using antisera directed against these peptides in immunological studies, several vasopressin and oxytocin homologues have been identified and isolated from several invertebrate sources. These include the vasopressin-like diuretic hormone from the migratory locust, *Locusta migratoria* (Proux *et al.*, 1987), and the conopressins from the gastropod molluscs, *Conus geographus, Conus striatus* (Cruz *et al.*, 1987), and *Lymnae stagnalis* (van Kesteren *et al.*, 1992). Oxytocin-like peptides which have been isolated are cephalotocin from the cephalopod mollusc, *Octopus vulgaris* (Reich, 1992), and annetocin from the earthworm, *Eisenia foetida* (Oumi *et al.*, 1994) (Table 1.2).

Table 1.2. Members of the vasopressin-oxytocin superfamily. Residues which are common to all the members are in bold face type (modified from Reich, 1992).

Vertebrate vasopressin related peptides[1]										
Cys	Tyr	Phe	Gln	**Asn**	**Cys**	**Pro**	Arg	**Gly-NH$_2$**	Arginine vasopressin	
Cys	Tyr	Phe	Gln	**Asn**	**Cys**	**Pro**	Lys	**Gly-NH$_2$**	Lysine vasopressin	
Cys	Phe	Phe	Gln	**Asn**	**Cys**	**Pro**	Arg	**Gly-NH$_2$**	Phenypressin	
Cys	Tyr	Ile	Gln	**Asn**	**Cys**	**Pro**	Arg	**Gly-NH$_2$**	Vasotocin	
Vertebrate oxytocin-related peptides[1]										
Cys	Tyr	Ile	Gln	**Asn**	**Cys**	**Pro**	Leu	**Gly-NH$_2$**	Oxytocin	
Cys	Tyr	Ile	Gln	**Asn**	**Cys**	**Pro**	Ile	**Gly-NH$_2$**	Mesotocin	
Cys	Tyr	Ile	Ser	**Asn**	**Cys**	**Pro**	Ile	**Gly-NH$_2$**	Isotocin	
Cys	Tyr	Ile	Ser	**Asn**	**Cys**	**Pro**	Gln	**Gly-NH$_2$**	Glumitocin	
Cys	Tyr	Ile	Gln	**Asn**	**Cys**	**Pro**	Val	**Gly-NH$_2$**	Valitocin	
Cys	Tyr	Ile	Asn	**Asn**	**Cys**	**Pro**	Leu	**Gly-NH$_2$**	Aspargtocin	
Invertebrate vasopressin-oxytocin-related peptides										
Cys	Leu	Ile	Thr	**Asn**	**Cys**	**Pro**	Arg	**Gly-NH$_2$**	Diuretic hormone	locust
Cys	Phe	Ile	Arg	**Asn**	**Cys**	**Pro**	Lys	**Gly-NH$_2$**	Lysine conopressin	snail
Cys	Ile	Ile	Arg	**Asn**	**Cys**	**Pro**	Arg	**Gly-NH$_2$**	Arginine conopressin	snail
Cys	Tyr	Phe	Arg	**Asn**	**Cys**	**Pro**	Ile	**Gly-NH$_2$**	Cephalotocin	octopus
Cys	Phe	Val	Arg	**Asn**	**Cys**	**Pro**	Thr	**Gly-NH$_2$**	Annetocin	earthworm

[1] Amino acid sequences of vertebrate peptides (taken from Archer and Chauvet, 1988).

1.5 Vertebrate vasopressin characteristics

Neuropeptides of the vasopressin-oxytocin family have been well characterized. Members of the vertebrate vasopressin-oxytocin peptide hormone family contain a ring of six amino acid residues closed by a disulphide bridge between two cysteine residues at positions 1 and 6, and a side chain of three amino acid residues. The hormones of this family differ from each other in amino acids in positions 3, 4, and 8, with the exception of phenypressin (Table 1.2) (Archer and Chauvet, 1988).

Vasopressin, a neurohypophysial hormone, is a product of post-translational modifications of a larger precursor molecule (preprohormone), which consists of the peptide hormone, its carrier protein, neurophysin, and a glycopeptide of as yet unknown function (Richter, 1983; Richter, 1987). Based on sequence analysis of cDNA clones, the bovine vasopressin preprohormone was found to consist of 166 amino acid residues with a molecular weight of 17,310 daltons, and the rat counterpart consists of 168 amino acid residues with a molecular weight of 17,826 daltons (Land *et al.*, 1982; Land *et al.*, 1983).

The gene structure of vasopressin is even more complex. By restriction mapping and nucleotide sequence analysis, the rat and bovine vasopressin genes were found to be about 2 kbp (kilobase pairs) long (Schmale *et al.*, 1983; Ruppert *et al.*, 1984). The principal functional domains of the hormone precursor are encoded by three distinct exons separated by two intervening introns (Fig. 1.2).

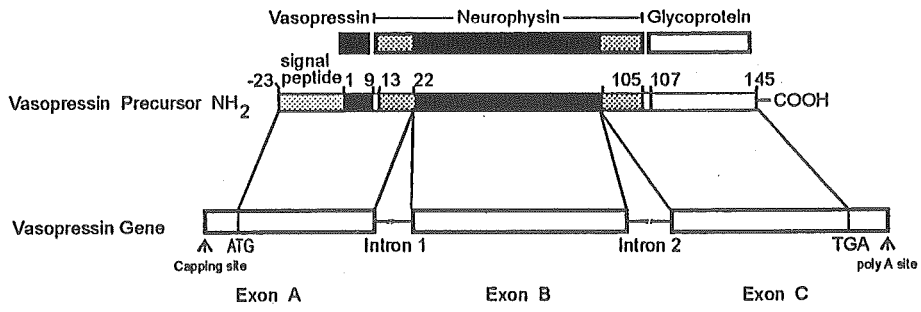


Figure 1.2. Structural organization of the rat vasopressin precursor and its gene (Schmale *et al.*, 1983).

The first exon (A) comprises the 5′-untranslated region, the putative signal peptide, the hormone vasopressin, and the N-terminus of neurophysin. The second exon (B) encodes the highly conserved centre part of neurophysin; and the third exon (C) encodes the remaining C-terminus of neurophysin, the glycoprotein, and the untranslated 3′ region with the polyadenylation site (Richter, 1987).

Though vertebrate-like neuropeptides of the vasopressin-oxytocin family have been isolated and identified from invertebrate sources, considerably less is known about the genes that encode these neuropeptides. van Kesteren *et al.* (1992) cloned and sequenced the cDNA encoding conopressin from the gastropod mollusc *Lymnaea stagnalis*. Their analysis of the deduced amino acid sequence for the *Lymnaea* preproconopressin indicates that it is organized much like the vasopressin-related precursors of the vertebrates. Preproconopressin consists of a signal peptide followed by conopressin, and a conserved neurophysin domain having a divergent glycoprotein-homologous C-terminal sequence.

1.6 The aims of this study

Based on previous work, isolation of the MIH has been hampered by the amount of material required to purify it, and also the lack of appropriate test to characterize its physiological activity both *in vivo* and *in vitro*. For example, 2000 sinus glands (SG) were required to purify a peptide (25 ng/SG) with MIH-like activity (Huberman and Aguilar, 1989), In addition, findings of both peptide and non-peptide compounds with MIH-like activity, and in the case of the peptides, with different molecular weights as well as amino acid sequences, have further indicated that the definitive MIH, and its role in the moulting control is much more complex than at first proposed.

In this study, I sought to isolate the putative MIH gene sequence from the lobster, using a different approach than has been reported. Findings of vasopressin homologues in other invertebrates, as well as the results of Van Herp and Bellon-Humbert (1982) and Mattson and Spaziani (1985a), prompted the use of the cloned

rat vasopressin gene as a probe to isolate the putative MIH gene sequence from the red rock lobster, *Jasus edwardsii*.

With this approach, I sought to address a few questions raised by results from other studies.

(a) Based on Mattson and Spaziani's (1985a) findings, would it be possible to isolate the putative MIH gene sequence from the lobster using sequence information that is known about the rat vasopressin gene.

(b) If gene sequences could be isolated by this method, are these sequences related to other vasopressin homologues isolated from other invertebrates, and to the published putative MIH sequences?

The specific aims of this study were:

1. To isolate the putative MIH gene sequences from the lobster *Jasus edwardsii*, using the rat vasopressin gene as a probe.

2. To characterize and compare the putative MIH cDNA of *Jasus edwardsii* with nucleotide sequences of other crustacean putative MIH sequences.

3. To examine tissue specific expression of these genes in *Jasus edwardsii*.

1.7 Experimental strategy

Several approaches were taken to isolated the putative MIH gene sequence from the lobster (Fig. 1.3). The presence of vasopressin-like gene sequences in the lobster was initially tested by Southern blot analysis using the rat vasopressin gene as a probe. However, subsequent screening of the lobster genomic DNA library with the rat vasopressin gene suggested that a more specific probe was required. PCR using primers based on the rat vasopressin gene was used to amplify the putative MIH gene sequence(s) in the lobster. The specificity of the probe was then demonstrated

by *in situ* hybridization to the lobster eyestalk and northern blot analysis. Based on these observations and sequence analysis, the PCR product was shown to contain coding sequences. This PCR product was subsequently used as a probe to isolate the putative MIH gene sequence from a cDNA library constructed from the lobster eyestalk, and also from the lobster genomic DNA library.

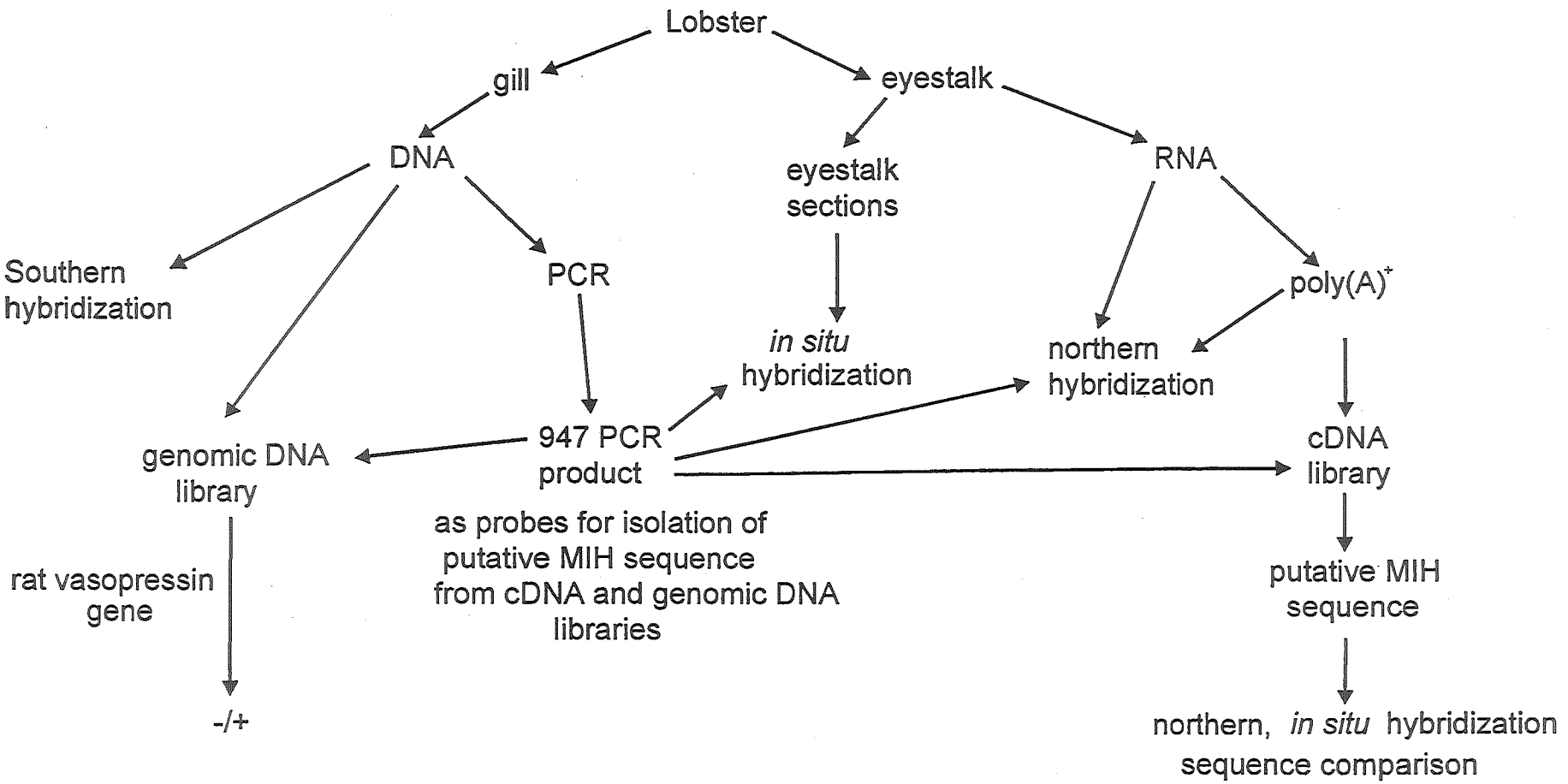


Figure 1.3. Diagram illustrating the experimental strategy used for isolating the putative MIH gene sequence

CHAPTER II

Search for a putative moult-inhibiting hormone (MIH) gene from the lobster *Jasus edwardsii*

2.1 INTRODUCTION

As reviewed in Chapter I, in decapod crustaceans, the X-organ/sinus gland complex located in the eyestalks is the source of several neuropeptide regulatory factors responsible for a variety of physiological effects (Quackenbush, 1986). Among these are the members of a peptide family which consists of the crustacean hyperglycemic hormone (CHH), a putative moult-inhibiting hormone (MIH), and the vitellogenesis-inhibiting hormone (VIH) (reviewed in Keller, 1992). The crustacean MIH was first identified as a peptide by Rangarao (1965) who isolated a compound with moult-inhibiting effect from the crab *Ocypode macrocera*. Since then, the search for the putative MIH has been conducted by several research groups; polypeptides with moult-inhibiting activity have been isolated from several crustacean species with molecular weights varying between 1000 to 9000 (M_r).

Mattson and Spaziani (1985a) have presented immunological evidence which suggests that the MIH is structurally related to the vertebrate vasopressins. Members of the vasopressin-oxytocin family mimick MIH-like activity by suppressing the production of ecdysone by cultured Y-organs. Both lysine vasopressin (LVP) and MIH (sinus gland extracts) have similar affinity characteristics with antibody to arginine vasopressin (AVP).

Based on these observations, I sought to isolate the putative MIH gene sequence from the lobster *Jasus edwardsii*. Using PCR primers derived from the rat vasopressin gene, a 947 bp fragment was amplified from lobster DNA. Sequence

analysis of the 947 bp fragment suggested that it consisted of both coding and non-coding regions. *In situ* hybridization of this fragment to lobster eyestalk sections showed tissue specific hybridization of this probe to the neurosecretory regions of the eyestalk. Preliminary observations suggested that the expression of this probe was related to the moulting cycle. The 947 bp fragment also hybridized to two mRNA bands, 0.68 and 1.68 kb in size, isolated from the eyestalk in northern blot analysis, which suggested that the 947 bp PCR product could contain sequences partially representing the gene sequence encoding the putative MIH from the lobster *Jasus edwardsii*.

2.2 MATERIALS AND METHODS

2.2.1 Detection of the vasopressin-like gene sequence by Southern hybridization

2.2.1.1 Isolation of lobster DNA

Live lobsters, *Jasus edwardsii*, were obtained from the Kaikoura coast in New Zealand. High molecular weight DNA was isolated from the lobster according to a method adapted from McGinnis *et al.* (1983). Gill tissue (8-13 g) was dissected from a lobster which had been chilled on ice for at least 30 min. The fresh tissue was cut into small pieces, snap frozen in liquid nitrogen, and ground to a fine powder in liquid nitrogen in a precooled mortar and pestle at -70°C. The powdered tissue was suspended in 10 volumes of ice cold homogenization buffer (100 mM NaCl, 30 mM Tris-HCl, pH 8.0, 10 mM EDTA, 10 mM mercaptoethanol, 0.5% Triton X-100), and homogenized on ice in a Dounce tissue grinder using pestle A. The homogenate was then filtered through a nylon filter (mesh, 80 μm), and centrifuged at 6700 x *g* for 10 min at 4°C, to pellet the nuclei.

The nuclei were washed by resuspending in extraction buffer (100 mM Tris-HCl, pH 8.0, 100 mM NaCl, 20 mM EDTA), at 1/2 the homogenate volume, and pelleted as before. The nuclei were then homogenized gently in 1v:1w of original tissue, in extraction buffer, and proteinase K and SDS were added to concentrations

of 100 µg/ml and 1% (w/v), respectively. The solution was mixed gently after the addition of each component, and incubated at 50°C for 1 h.

Proteins were then removed by extraction with equal volumes of phenol saturated with TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0), and chloroform (24:1, chloroform : isoamyl alcohol). The two phases were mixed thoroughly by gentle inversion, followed by centrifugation at 15,800 x *g* for 20 min at 15°C to separate the aqueous and organic phases. The aqueous phase was transferred to a fresh tube using a large-bore Pasteur pipette, and the interphase was extracted with TE as the two phases were difficult to separate. The aqueous phases were combined and extracted with phenol-chloroform, and then twice with chloroform.

RNA was removed by digestion with 100 µg/ml pancreatic RNase A (Boehringer Mannheim) (heat treated to inactivate any DNase activity), at 40°C for 1 h. The mixture was then extracted twice with equal volumes of phenol and chloroform, and twice with chloroform. The aqueous phase containing DNA was concentrated by extraction with 2-butanol; an equal volume of 2-butanol was added to the aqueous phase, and the two phases were mixed and centrifuged briefly to separate them. The upper organic phase was removed and the procedure repeated until the required volume was obtained. The aqueous phase was then dialysed against a 200-fold excess of TE, pH 8.0, at 4°C for several hours to remove traces of contaminants, such as phenol and excess salts.

The concentration of DNA was estimated using spectrophotometry, where one absorbance unit at 260 nm (A_{260}) is equivalent to 50 µg/ml of double stranded DNA (Sambrook *et al.*, 1989). The concentration of lobster genomic DNA isolated using this procedure was generally around 200-300 µg/ml and the A_{260}/A_{280} ratio was between 1.76-2.0. The ratio suggested that the DNA preparation could contain traces of RNA, but was free of phenol and protein contaminants (Sambrook *et al.*, 1989). The yield of DNA per gram of tissue varied for different tissues; approximately 200 µg DNA/g lobster gill tissue, and 800 µg DNA/g of rat liver tissue were isolated, using this method. DNA was stored at 4°C until required.

2.2.1.2 Southern hybridization

The presence of vasopressin-like gene sequences in the lobster was detected by Southern hybridization to probe pBS 13+, a Bluescript m13+ plasmid containing the rat vasopressin gene, obtained from Dr. Evita Mohr (Universitätskrankenhaus Eppendorf, Hamburg, Germany).

Twenty micrograms of lobster and rat genomic DNA, at concentrations of 200 µg/ml, were digested with 6 U and 5 U/µg, respectively, of the following restriction enzymes: *Bgl*II, *Eco*RI, *Hin*dIII and *Pst*I (Boehringer Mannheim). Digestion was carried out at 37°C for 16 h. An equal amount of restriction enzyme was added to each lobster sample and digestion carried out for a further 2 h. The digested DNA was ethanol precipitated at −20°C overnight, pelleted by centrifugation at 12,000 x *g*, for 20 min at 4°C, washed in 70% ethanol, air dried and dissolved in TE (0.6 µg/µl). Samples were run in sample buffer III (0.04% bromophenol blue, 0.04% xylene cyanol, 5% glycerol; Sambrook *et al.*, 1989) in a 1% agarose gel in TBE (89 mM Tris-borate, 2 mM EDTA, pH 8.0) with ethidium bromide (0.5 µg/ml), at 0.85 V/cm for 18 h.

DNA was then transferred by alkaline transfer (Amersham) to nylon membrane (Zeta-Probe, Bio-Rad). The gel was incubated in 0.25 M HCl for 2 x 15 min, and in 1.5 M NaCl/0.5 M NaOH for 30 min. DNA was transferred in a solution of 1.5 M NaCl/0.25 M NaOH by capillary transfer according to Sambrook *et al.* (1989) for 16-24 h. The membrane was washed for 10 min in 2 x SSC (0.3 M NaCl, 30 mM trisodium citrate), dried at 60°C, and DNA cross-linked to the membrane by exposure to 254 nm ultraviolet light at a distance of 6 cm for 1 min.

The membrane was prehybridized in 50% formamide, 0.25 M $NaHPO_4$ (pH 7.2), 7% (w/v) SDS, 1 mM EDTA (150 µl soln./cm^2) at 43°C for at least 1 h in a shaking water bath. Hybridization was carried out using 250 ng of heat denatured [α-^{32}P] dCTP radioactively labelled probe (10^6 cpm/µg) in a 5-10 ml volume, for 16-24 h. After hybridization, the membrane was rinsed briefly in 2 x SSC, followed by successive washes in 2 x SSC/0.1% SDS and 0.5 x SSC/0.1% SDS for 15 min at

ambient temperature, with a final wash in 0.5 x SSC/0.1% SDS at 65°C for 15 min or 0.1 x SSC/0.1% SDS at 65°C for 15 min.

Southern blots were stripped of probe by washing the membranes in 0.1 x SSC/ 0.5% SDS for 2 x 20 min at 95°C. The membranes were checked by overnight exposure to X-ray film, and stored at −20°C for several weeks before reprobing.

2.2.1.3 Radiolabelling of probes

DNA probes used in Southern hybridization were prepared by nick translation according to Sambrook *et al.* (1989). Approximately 250 ng of pBS 13+, containing the rat vasopressin gene was nick translated in a 25 µl reaction using [α-^{32}P] dCTP (25 µCi, 3000 Ci/mmol, Amersham), at 16°C for 1 h. The radiolabelled DNA was separated from the unincorporated dCTPs by chromatography on a P-60 sephadex column, followed by precipitation in 0.3 M NaCl and 2 volumes of ethanol at -80°C for 1 h, or precipitation in 2 M ammonium acetate (pH 7.5) and 1 volume of isopropanol at room temperature for 10 min. The probe was then pelleted by centrifugation at 16,000 x *g* for 15 min, washed in 70% ethanol, dried under vacuum and dissolved in 100 µl TE. The specific activity obtained was approximately 10^6-10^7 cpm/µg DNA. Radioactivity was measured for a standard volume of 100 µl in an Eppendorf, in a Quick-count benchtop radioisotope counter (BIOSCAN/QC.4000XER).

2.2.2 Amplification of the vasopressin-like gene sequence by polymerase chain reaction (PCR).

2.2.2.1 PCR

Since the experiments of Mattson and Spaziani (1985a) suggested that the crustacean MIH was immunologically related to mammalian vasopressin, primers were designed to amplify by PCR the vasopressin-like gene sequence in the lobster *Jasus edwardsii*. The two oligonucleotides used for PCR were:

Primer 1, 5′-AACTGCCCAAGAGGAGGC-3′; and

Primer 2, 5′-AGGGCCGCAGGGGAGACA-3′.

The primers were derived from the rat vasopressin gene sequence (Schmale *et al.*, 1983). Primer 1 was based on the DNA sequence coding for amino acid residues 5 to 10. The amino acid residues 5 to 9, are common to most of the peptides in the vasopressin-oxytocin family. Amino acid residue 10 is the glycine residue which is involved in the maturation process leading to the active hormone (Archer, 1981; Richter, 1987). Primer 2 was the DNA sequence of the highly conserved part of exon B in the rat vasopressin gene. These two primers would amplify a 1239 bp fragment from the rat vasopressin gene (Fig. 2.1).

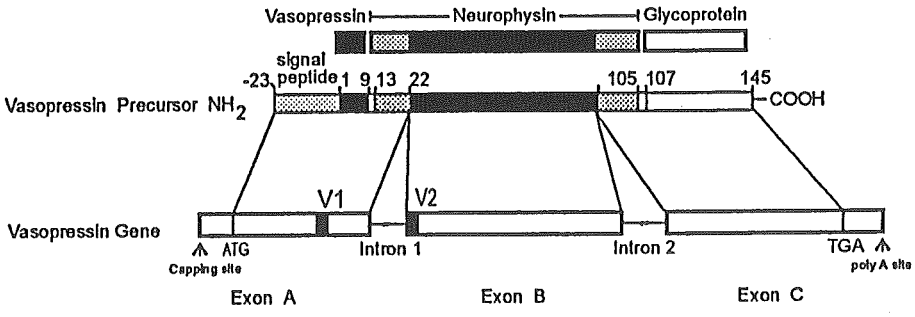


Figure 2.1 Structural organization of the rat vasopressin precursor and its gene (Schmale *et al.*, 1983). The location of the primers used in PCR are indicated (V1, V2).

The initial PCR was carried out on lobster DNA with varying $MgCl_2$ concentrations (0.5-5.0 mM) and annealing temperatures (45°, 50° and 55°C) in a 50 µl reaction volume containing 0.8 µM each of primers 1 and 2, 200 µM dNTPs, 50 mM KCl, 10 mM Tris-HCl (pH 9.0), 0.1% Triton X-100, and 1 U *Taq* DNA polymerase (Promega). Five hundred nanograms of lobster DNA was used in each reaction, and as positive controls, 500 ng of rat DNA and 10 ng of the plasmid containing the rat vasopressin gene were used. A water control was run with each set of reactions to check for contamination. PCR was carried out in the Autogene II thermal cycler (Grant) or Perkin Elmer DNA thermal cycler 480. PCR was carried out for 30 cycles with a denaturation step at 94°C for 1 min, an annealing step at 55°C for 1 min, and an extension step at 72°C for 1 min, with a final extension step of 7 min. Ten microliters of PCR products were analyzed by electrophoresis on a

1.6% agarose gel, and visualized by staining in ethidium bromide and viewed over a UV transilluminator.

2.2.2.2 Re-amplification of PCR products

The DNA from two major PCR products, 564 bp and 947 bp, was recovered according to a method described by Koenan (1989). These two major PCR products were re-amplified, and gel purified using the Prep-A-Gene kit from Bio-Rad. The 947 bp purified PCR fragment was subsequently used for the screening of cDNA and genomic DNA libraries (Chapters III).

2.2.3 Characterization of the PCR products

To determine whether the PCR products contained sequences related to the putative MIH gene sequence, the two major PCR products, 564 bp and 947 bp, were used as probes in Southern blot analysis of lobster and rat genomic DNA, *in situ* hybridization to lobster eyestalks and northern blot analysis. If either of the PCR products gave results which were consistent with what one would expect for the putative MIH gene sequence, such as expression in the eyestalk X-organs-sinus gland complex, then the PCR product would be a likely candidate as a probe for isolating the putative MIH gene sequence from both a cDNA and lobster genomic DNA library.

2.2.3.1 Detection of the PCR products in the lobster and rat genome

The 564 bp and 947 bp PCR products were used as probes in DNA dot blot analysis and Southern hybridization to lobster and rat DNA to determine whether these two PCR products contained DNA sequences that were specific to the lobster, or found in both the rat and lobster. Hybridization of the probes, using high stringency conditions, to distinct bands only within the lobster digests, would indicate that the amplified products contained DNA sequences specific to the lobster, and could represent part of a gene sequence.

The PCR products (564 bp and 947 bp fragments) were radiolabelled by the random priming reaction using either the NEBlot kit (Biolab), or the random priming

kit from Boehringer Mannheim. Between 10 to 100 ng of DNA was used in each reaction using [α-^{32}P] dCTP (50 µCi, 3000 Ci/mmol), and the probes were purified by separation on a P-60 sephadex column (section 2.2.1.3). The specific activity obtained for random primed probes was between 10^8-10^9 cpm/µg DNA.

For DNA dot blot hybridization, 1 µl (100 ng) of lobster DNA and rat DNA was spotted on nylon membrane (Hybond-N, Amersham). DNA was denatured by wetting the membrane in 1.5 M NaCl, 0.5 M NaOH for 7 min. The membrane was then neutralized in 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.2), 1 mM EDTA for 6 min with one solution change, and then washed briefly in 2 x SSC. The membrane was air dried and the DNA cross-linked on the membrane by UV illumination.

Hybridization was essentially the same as described for Southern hybridization (section 2.2.1.2), except that the formamide prehybridization solution was replaced with an aqueous solution containing 6 x SSC, 5 x Denhardt's solution (0.02% (w/v) of bovine serum albumin (BSA), Ficoll and polyvinylpyrrolidone (PVP)), 0.5% (w/v) SDS, and 0.5% (w/v) blocking agent (Amersham) at 65°C. Post hybridization washes were: 2 x SSC for 2 x 15 min at 65°C, 2 x SSC/0.1% SDS for 30 min at 65°C, and a final wash of 0.5 x SSC/0.1% SDS for 10 min at 65°C.

Southern blot analysis of lobster and rat genomic DNA, previously probed with the rat vasopressin gene, was carried out using the 564 bp and 947 bp PCR products as probes.

2.2.3.2 *In situ* hybridization of lobster eyestalk sections using the 564 bp and 947 bp PCR fragments

It has been previously established that the crustacean MIH is produced in the X-organs/sinus gland complex of the eyestalk (reviewed in Lachaise *et al.*, 1993). Therefore, *in situ* hybridization of the PCR products to the neurosecretory regions of the lobster eyestalk would suggest that the expression of the PCR products is tissue specific, and the PCR products may be related to the crustacean MIH gene, or even other crustacean eyestalk neuropeptide genes.

The *in situ* hybridization methods used were modified from previous studies (Bloch *et al.*, 1986; Chesselet, 1990; Leitch *et al.*, 1994; Pardue, 1985; Zeller *et al.*, 1989). RNase contamination was avoided by wearing gloves. Plasticware, glassware, and acid washed glass slides were baked at 160°C for 4 h. The solutions used were treated with 0.2% diethyl pyrocarbonate (DEPC), incubated at 37°C overnight and autoclaved to remove traces of DEPC.

Animals were maintained in seawater at 15°C, on a diet of mussel, for up to several months. Eyestalks were removed from a lobster which had been previously chilled on ice for approximately 40 min. The overlying cuticle and muscle were dissected from the eyestalks in ice cold crustacean saline (224 mM NaCl, 7.51 mM KCl, 12.5 mM $CaCl_2$, 0.49 mM $MgCl_2$, 0.97 mM $MgSO_4$, 4.44 mM H_3BO_3 adjusted to pH 7.5 with NaOH; Midsukami, 1979). The dissected eyestalks were fixed in freshly prepared 4% paraformaldehyde at 4°C for 50-60 min, washed twice in phosphate-buffered saline (PBS), and then left overnight in 15% sucrose in PBS at 4°C. Phosphate-buffered saline was made by mixing 130 mM NaCl, 10 mM Na_2HPO_4; and 130 mM NaCl, 10 mM NaH_2PO_4 to obtain pH 7.2 (Zeller *et al.*, 1989).

Paraformaldehyde fixative was prepared by dissolving 4% paraformaldehyde in double distilled water (ddH_2O) (2/3 of the final volume) at 60°C with several drops of 2N NaOH. The solution was made up to volume with 3 x PBS, adjusted to pH 7.2 with 1 M HCl and filter sterilized (Zeller *et al.*, 1989).

The tissue was mounted in O.C.T. compound (Tissue-Tek) and snap frozen in liquid nitrogen vapours. This alternative method was used since the recommended process of snap freezing tissue blocks, by plunging into isopentane cooled in liquid nitrogen, caused tissue blocks to shatter occasionally. Sixteen micron sections were cut at −26°C, in a Starlet *2212* Cryostat. Cryostat cut sections were collected on warm gelatin-chrome alum-subbed slides, dried for 2 min on a hot plate (60°C), and stored at −80°C. Subbed slides were prepared by dipping slides into subbing solution

(0.5% gelatin (bloom 300), 0.5% chrome alum) for 2 min. The slides were air dried overnight in a dust free environment and stored at -80°C (Bloch *et al.*, 1986).

Frozen slides were brought to room temperature, and then soaked in prehybridization buffer (4 x SSC, 1 x Denhardt's) for 10 min and 45 min at room temperature. Following this, the slides were soaked in 4 x SSC for 2 x 10 min at room temperature, and then for 10 min at room temperature in a freshly prepared solution of 4 x SSC, 0.1 M triethanolamine, and 0.5% acetic anhydride. The sections were then rinsed briefly in 70% ethanol, twice in absolute ethanol, and air dried.

Probes were labelled with [α-^{32}P] dCTP (50 µCi, 3000 Ci/mmol) to 10^8 cpm/µg of DNA by random priming (section 2.2.3.1). After purification by column chromatography and ethanol precipitation, probes were dissolved in 100 µl of hybridization buffer (100 µl 8 x SSC, 100 µl deionized formamide, 10 µl 20% Sarkosyl, 1 x Denhardt's, 20 µl 1.2 M phosphate buffer, pH 7.2).

The radioactivity was estimated using a BIOSCAN/QC.4000XER benchtop radioisotope counter. The probe was denatured by boiling for 8 min and quenching on ice, and then made up to the appropriate volume for the number of slides used. Approximately 500,000 cpm or 1-5 ng of probe was applied on each slide in a volume of 20 µl per 20 x 22 mm area, and tissue sections were overlaid with acetone washed, siliconized (Sigmacote) glass cover slips. Hybridization was carried out for 16 h at 42°C in a large plastic container lined with paper towels dampened with prehybridization buffer.

Following hybridization, the coverslips were washed off in 4 x SSC, and slides were soaked in 2 x SSC for 10 min and 45 min at room temperature. The slides were then soaked in 1 x SSC at 40°C for 10 min, then in fresh SSC for 45 min. The slides were dehydrated in 70% ethanol and then in two changes of absolute ethanol. Air dried sections were exposed to X-ray film for 0.5-16 h, to estimate the required exposure time to liquid emulsion. The slides were then coated with Ilford LM-1 liquid emulsion (diluted with 0.5 volumes of distilled water) at 43°C, air dried for at

least 1 h, and stored in the dark at 4°C with desiccant for 1-4 days (approximately 28 times the exposure to X-ray film). After developing the slides with Kodak D-19 developer at 15-19°C, the sections were stained in Ehrlich's haemotoxylin stain (Merck) for 30 min, followed by three quick dips in 0.5% conc. HCl, and then gently washed in running water for at least 10 min. Sections were then rinsed in xylene for 2 min, mounted in Eukitt (Fisher Scientific), and dried at 37°C for 2 days.

2.2.3.3 Northern hybridization of total RNA and mRNA to the 947 bp PCR fragment

To determine whether the PCR fragments contain DNA sequences coding for RNA, and to determine the size of these mRNA transcripts, northern hybridization was carried out on total RNA and mRNA isolated from several different tissues of the lobster using the 947 bp fragment as a probe.

Solutions used in total RNA and mRNA extraction were treated with 0.2% DEPC at 37°C overnight, then autoclaved to inactivate the remaining DEPC. Glassware and plasticware used in the extraction procedure were either baked at 160°C for 4h or, soaked in a 0.2% solution of DEPC at 37°C overnight and then rinsed with distilled water before autoclaving.

(i) Extraction of total RNA

Eyestalks were ablated from live lobsters obtained from the Pacifica Kaikoura Ltd. Fisheries at Kaikoura, transported on dry ice to the laboratory, and stored in liquid nitrogen (approximately 10 months). Total RNA was extracted from lobster eyestalks by the acid guanidinium thiocyanate-phenol-chloroform extraction procedure of Chomczynski and Sacchi (1987). The following modifications were made to the procedure; the cuticle was dissected off slightly thawing eyestalks and the remaining soft tissue was macerated in denaturing solution (0.12 g per eyestalk, 0.26 g per ml denaturing solution) with a Tissuemizer (Tekmar) for 10 s at room temperature. Purified RNA was dissolved in DEPC-treated ddH_2O.

The concentration of RNA was estimated at 3.4 µg/µl by spectrophotometry readings at A_{260}; where 40 µg/ml RNA is equivalent to 1 A_{260} (Sambrook *et al.*, 1989), and 1.5 µg/µl by spotting on a 1% agarose gel (0.5 µg/ml of ethidium bromide) against known amounts of tRNA. Purple pigments in the tissue could have interfered with the spectrophotometry readings, thus giving a higher value. The yield of total RNA isolated was estimated at 0.38 mg/g wet tissue, or 45 µg/lobster eyestalk.

RNA was stored at −80°C, in small aliquots, with 3 volumes of absolute ethanol, and recovered by adding 0.1 volume of 3 M sodium acetate, pH 5.2, and centrifuging at 12,000 x *g* for 20 min at 4°C. The RNA was washed in 70% ethanol, air dried, and dissolved in DEPC-treated ddH_2O at concentrations of 0.5-10 µg/µl.

Total RNA was also extracted from the epithelial tissue, gill, heart, hepatopancreas, and abdominal flexor muscle from one individual, for nothern blot analysis, using the same procedure, but quantities were adjusted for each tissue type. The yield of total RNA per g of wet tissue was estimated at 0.83 mg/g epithelial tissue, 0.36 mg/g gill tissue, 0.67 mg/g of heart tissue, 3.15 mg/g hepatopancreas, and 0.42 mg/g muscle tissue, as determined by spectrophotometry readings at A_{260}.

(ii) Purification of mRNA (polyadenylated RNA)

Polyadenylated RNA ($poly(A)^+$ RNA) was extracted from total RNA using the polyATtract mRNA isolation system from Promega. This system used a biotinylated oligo(dT) primer which hybridized at high efficiency in solution to the 3′ poly(A) region in mature eukaryotic messenger RNA. The hybrids were then captured and washed at high stringency using streptavidin coupled to paramagnetic particles, and a magnetic separation stand. The mRNA was eluted from the solid phase by the addition of ribonuclease-free deionized water. The amount of total mRNA isolated was estimated by spectrophotometry at A_{260} of the eluted solution, using cuvettes that had been previously soaked in concentrated HCl:methanol (1:1) for 1 h. The yield obtained was approximately 1-1.5% of the total RNA used (approximately 1 mg). Polyadenylated RNA was stored in 3 volumes of ethanol at −80°C, and recovered by

the addition of 0.1 volume of 3 M sodium acetate, pH 5.2, and centrifugation at 12,000 x *g*, for 20 min at 4°C. The mRNA isolated by this technique was used for northern hybridization, cell free translation, and cDNA synthesis (see Chapter III).

(iii) Northern hybridization of total RNA and mRNA to the 947 PCR fragment

Total RNA and mRNA isolated from lobster eyestalks, epithelial tissue, gill, heart, hepatopancreas and muscle tissue were separated electrophoretically on a 1.2% agarose gel. RNA was dissolved in 4.5 µl DEPC-treated water (between 0.3-10 µg/µl) and sample treatment buffer added (2 µl of 5 x TBE, 3.5 µl of 37% formaldehyde, 10 µl of deionized formamide; Sambrook *et al.*, 1989). One microlitre of ethidium bromide (1 µg/ml) was added to each sample for visualization of the RNA on the agarose gel (Rosen and Villa-Komaroff, 1990). Samples were heated at 65°C for 15 min, and chilled on ice. One tenth volume of glycerol dye loading buffer (0.2 M EDTA, 50% glycerol, 0.25% bromophenol blue, 0.25% xylene cyanol) was added to each sample before loading on the agarose gel. DNA molecular weight markers (*Hin*dIII digested λDNA) were denatured by heating at 100°C for 10 min, and loaded immediately before electrophoresis commenced.

Agarose gels were run at 120 V (4V/cm) in 0.5 x TBE for 1-3 h depending on the size of the gel. Gels were run without formaldehyde as the denaturation treatment of RNA molecules prior to loading was sufficient to maintain RNA molecules in the denatured state for at least 3 h (Liu and Chou, 1990). The amount of total RNA (25-28 µg) and poly(A)$^+$ RNA varied in the amount available for electrophoresis: epithelial tissue (26 µg total and 1.0 µg poly(A)$^+$), eyestalk (25 µg total, 1.44 µg poly(A)$^+$), gill (25 µg total, 2 µg poly(A)$^+$) heart (0 µg total RNA, 1.35 µg of poly(A)$^+$ RNA), hepatopancreas (28 µg total, 2.7 µg poly(A)$^+$) and muscle (25 µg total and 2.43 µg poly(A)$^+$).

RNA was transferred by capillary action in 10 x SSC on to nylon membrane (Hybond-N, Amersham), for 16 h according to Sambrook *et al.*, (1989). Following transfer, the membrane was air dried and the RNA crosslinked to the membrane by

UV transillumination. The membrane was then rinsed briefly in 2 x SSC, to wash off excess salt, prior to hybridization.

Northern hybridization was carried out according to the method described by Lee *et al.* (1992). Membranes were prehybridized at 150 µl/cm^2 in prehybridization solution (0.43 M sodium phosphate, pH 7.2, 7% SDS, 20 mM EDTA, 1% BSA), at 60°C with shaking for at least 1 h. The membranes were then hybridized in 10-20 ml of prehybridization solution containing 25 ng of denatured random prime radiolabelled probe (section 2.2.3.1) at 60°C for 16 h. Post-hybridization washes carried out were: 2 x SSC/0.1% SDS at 42°C for 2, 5, and 15 min, followed by 0.2 x SSC/0.1% SDS at 60°C for 3 x 15 min.

2.2.4 Sequence analysis of PCR products

The PCR products were partially sequenced from both ends using double stranded (ds) cycle sequencing. Based on the partial sequences, internal primers were constructed to amplify specifically the larger PCR product, to give a 850 bp fragment. To sequence further in, the PCR products were subcloned into pBluescript pBS M13+ using T-A cloning. The entire sequence was then obtained by a combination of ds sequencing using T7 DNA polymerase sequencing, and automated sequencing (Centre for Gene Research, University of Otago, Dunedin, New Zealand). The final sequence was confirmed by automated sequencing using an internal primer.

2.2.4.1 ds cycle sequencing of PCR products

The PCR products were partially sequenced from both ends using the ds cycle sequencing kit from BRL, in either an Autogene II or a Perkin Elmer DNA thermal cycler 480. Fifty femtomole of PCR product and 1 pmol of PCR primer were used in each sequencing reaction. The primers were 5′-end labelled with 10 µCi of [γ-^{33}P] ATP (>5000 Ci/mmol) using 1 U of polynucleotide kinase, at 37°C for 30 min, and the reaction terminated by heating at 55°C for 5 min.

Sequencing reactions were carried out as specified in the manufacturer's instructions, using the following parameters: 20 cycles of 30 s at 95°C, 30 s at 55°,

60 s at 70°C; followed by 10 cycles of 30 s at 95°C, 60 s at 70°C. The reactions were terminated by the addition of 5 µl of Stop Solution, and 1.5-2 µl of each reaction was analyzed on a 6% polyacrylamide-8.32 M urea sequencing gel in TBE (89 mM Tris-borate, 2.5 mM EDTA, pH 8.3), using a S-2 BRL sequencing gel appparatus with 43 x 30 cm glass plates. Electrophoresis was carried out at 1600-1700 V, with a constant power set at 60-70 W, and current <45 mA, using a 2103 Power Supply (LKB).

To read sequences of 250 to 350 nt, samples were run for approximately 2 h (short run), and 4 h (long run). Samples were loaded at the beginning of the run (long run), and 20 min after the bromophenol blue dye front had migrated off the edge of the gel, samples were loaded for the short run. Gel electrophoresis was terminated when the bromophenol blue dye front of the short run had just come out of the gel (approximately 4 h). The gel was transferred on to a piece of filter paper (Bio-Rad), and dried under vacuum at 80°C for 30 min (sequencing cycle) on a gel dryer (Bio-Rad). The dried gel was exposed to X-ray film (Hyperfilm, Amersham or Kodak X-Omat AR film), with intensifying screens at -80°C for 2-7 days.

2.2.4.2 Sequencing of PCR products cloned into pBS m13+

Since ds cycle sequencing would allow accurate readings of 250-280 nt, both the 564 bp and 947 bp PCR products were cloned into a plasmid vector, pBluescript (pBS m13+) by T-A cloning, for sequencing of these products.

PCR products were subcloned into the *Sma* I restriction site of pBS m13+ using T-A cloning (Clark, 1988; Marchuk *et al.*, 1991). The plasmid vector was cut with *Sma* I, and then purified with phenol-chloroform, ethanol precipitated, and the pellet resuspended in TE (pH 8.0). Deoxythymidine 5′-triphosphate (T) overhangs were generated by incubating 5 µg of blunt-ended vector DNA with 5 mM dTTP, 1.5 mM $MgCl_2$, 5 U *Taq* DNA polymerase, in 100 µl of PCR reaction buffer, at 75°C for 2 h (Finney, 1994). The DNA was later extracted with phenol-chloroform, recovered by ethanol precipitation, and dissolved in TE (pH 8) at a concentration of 0.2 µg/µl.

Ligation reactions of 1:6 molar quantities of vector (100 ng) to insert (110-180 ng) were set up according to Sambrook *et al.* (1989). Each 10 µl reaction contained DNA, 1 mM ATP, 1 U T4 DNA ligase (BRL), 50 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 20 mM DTT. Ligation was carried out for 16 h at 16°C, and the ligated DNA was then used to transform DH5α competent cells (Maniatis *et al.*, 1982). The entire mixture of transformed cells were then plated out on SOC plates (Appendix I) containing 50 µg/ml ampicillin, 20 µg/ml Xgal, and 0.5 mM IPTG, and grown at 37°C overnight.

Bacterial colonies carrying putative recombinant plasmids, which appeared white or pale blue after an hour at 4°C, were toothpicked on to fresh agar plates, and transferred on to nylon membrane (Hybond-N, Amersham), as described in Sambrook *et al.* (1989) and according to the manufacturer's instructions. Hybridization of the filters to the 947 bp PCR product was carried out as described for DNA dot blot analysis (section 2.2.3.1), with a final wash in 0.1 x SSC/0.1% SDS at 65°C for 10 min. The structure of the recombinant plasmids was then verified by restriction analysis and Southern hybridization of boiling preparations of plasmid DNA (Sambrook *et al.*, 1989; adapted from Holmes and Quigley, 1981).

Double stranded plasmid DNA was sequenced by the dideoxy sequencing method of Sanger *et al.* (1977), using the T7 Sequencing kit (Pharmacia LKB). Plasmids were sequenced from both ends using 35 ng of T3 and T7 promoter primer (Promega), 2 µg of plasmid DNA and [α-^{32}P] dCTP (10 µCi, 3000 Ci/mmol), and 2 µl of each reaction was analyzed on 6% polyacrylamide-8.32 M urea sequencing gels in TBE. Dried gels were exposed for autoradiography, without intensifying screens, for 6-18 h at -80°C. The remainder of the sequence data was obtained by automated sequencing of the templates (Centre for Gene Research).

2.2.4.3 Sequence analysis of PCR products

The DNA sequences were analysed using HIBIO DNASIS software system (Pharmacia LKB Biotechnology AB), to examine for open reading frames, predicted protein coding regions based on the theory developed by Fickett (1982), and to look

for DNA sequence similarity to other sequences present in the EMBL and GenBank nucleotide databases.

The DNA sequence was also translated into putative proteins in six reading frames for protein sequence similarity search using the Basic Local Alignment Search Tool (BLAST) developed by NCBI (Altschul *et al.*, 1990), which accessed the major nucleotide and protein sequence databases.

2.3 RESULTS

2.3.1 Southern hybridization of lobster DNA to the rat vasopressin gene

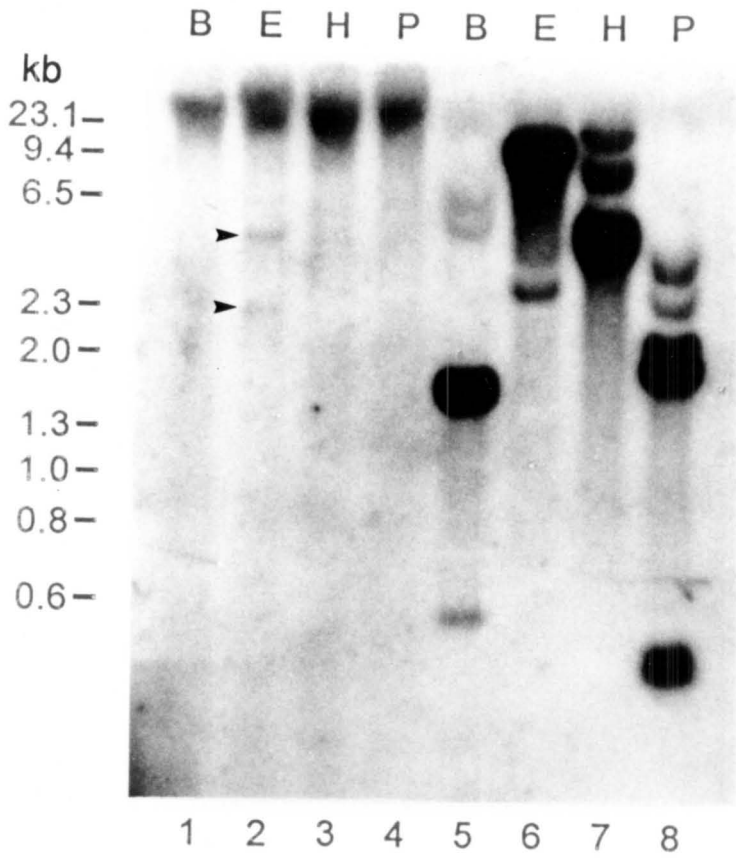
The presence of vasopressin-like gene sequences in the lobster was detected by Southern hybridization of restriction digests of lobster DNA using the rat vasopressin probe (Fig. 2.2). Restriction digests of rat DNA, and plasmid DNA were included as positive controls. The rat vasopressin gene hybridized to two bands at 4.5 and 2.3 kb in the *Eco*RI digest of lobster, *J. edwardsii*, genomic DNA, and also to the undigested high molecular weight lobster DNA in all four lanes (Fig. 2.2, lanes 1-4). The rat vasopressin gene also hybridized to several bands on the restriction digests of rat genomic DNA, with strong hybridization signals at approximately 6.1, 4.8, 1.6 and 0.5 kb (*Bgl*II), 9.4 and 2.8 kb (*Eco*RI), 14.7, 7.8 and 4.8 kb (*Hin*dIII), and 3.1, 2.4, 2.0, 1.8 and 0.4 kb (*Pst*I) (Fig. 2.2, lanes 5-8). In addition, the probe hybridized to the positive control, 10 ng of plasmid pBS 13+ (data not shown).

In a subsequent experiment, when a high stringency post-hybridization wash was included, hybridization of the rat vasopressin probe to undigested lobster DNA was significantly reduced (Fig. 2.3, lanes 1-3). The probe remained bound to the two *Eco*RI bands (4.5 and 2.3 kb), giving a stronger signal with the 2.3 kb band (Fig. 2.3, lane 2). Detection of vasopressin-like gene sequences in the lobster even at high stringency conditions indicated that the vasopressin-like gene sequence in the lobster had high homology to the rat vasopressin gene supporting the earlier findings of

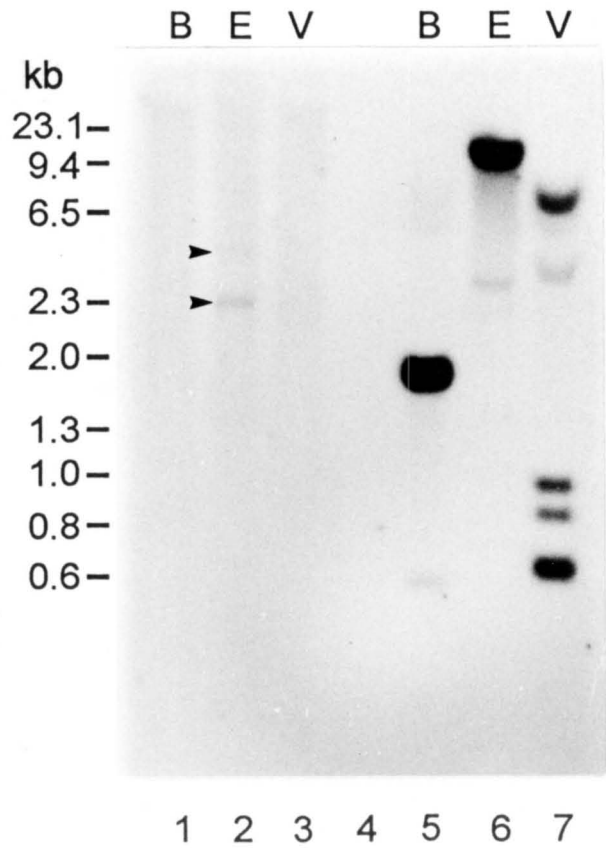
Figure 2.2. Southern blot analysis of genomic DNA from the lobster (lanes 1-4) and rat (lanes 5-8). Twenty micrograms of genomic DNA was digested with *Bgl*II (B), *Eco*RI (E), *Hin*dIII (H), or *Pst*I (P). The blot was hybridized with the rat vasopressin gene DNA probe, pBS 13+, and a final post-hybridization wash in 0.5 x SSC/ 0.1% SDS at 65°C for 15 min was used. The autoradiogram was exposed for approximately 9 days, at -70°C. ▸, locations of the 2.3 kb and 4.5 kb bands. Location of DNA size markers (λ Drigest III) are indicated (kb): 23.1, 23,130 bp; 9.4, 9,416 bp; 6.5, 6,557 bp; 2.3, 2,332 bp; 2.0, 2,027 bp; 1.3, 1,353 bp; 1.0, 1,078 bp; 0.8, 872 bp; 0.6, 603 bp.

Figure 2.3. Southern hybridization of lobster genomic DNA (lanes 1-3) and rat genomic DNA (lanes 5-7) using the rat vasopressin gene DNA as probe. Twenty µg of lobster genomic DNA was digested with 8.1 U/µg *Bgl*II (B), 11.3 U/µg *Eco*RI (E), and 9.5 U/µg *Pvu*II (V). The rat genomic DNA was digested using 5 U/µg DNA for all three enzymes. The blot was washed in a final post-hybridization wash of 0.1 x SSC/0.1% SDS at 65°C for 15 min, and exposed for autoradiography at -70°C for approximately 9 days. ▸, locations of the 2.3 kb and 4.5 kb bands. The location of DNA size markers (λ Drigest III) are indicated (kb).

2.2


B E H P B E H P
kb
23.1
9.4
6.5
2.3
2.0
1.3
1.0
0.8
0.6
1 2 3 4 5 6 7 8

2.3


B E V B E V
kb
23.1
9.4
6.5
2.3
2.0
1.3
1.0
0.8
0.6
1 2 3 4 5 6 7

vasopressin and neurophysin-like peptides in the eyestalk of the shrimp (van Herp and Bellon-Humbert, 1982).

The hybridization signal was also considerably reduced for the 6.1 and 4.8 kb bands in the *Bgl*II digests of rat DNA, and 2.8 kb band in the *Eco*RI digest (Fig. 2.3, lanes 5 and 6). As an *Eco*RI restriction site is absent in the rat vasopressin gene (Schmale *et al.*, 1983), the bands giving weak hybridization signals were probably rat oxytocin gene sequences binding to the rat vasopressin probe. The rat vasopressin and oxytocin gene show considerable homology both in organization and nucleotide sequence (Ivell and Richter, 1984); there is 59.8% homology over a 778 bp overlap between the two genes.

2.3.2 Amplification of the putative MIH gene sequence in the lobster using PCR

According to the results of Mattson and Spaziani (1985a), crustacean MIH was found to be structurally related to the vasopressins. Based on their findings, PCR primers, derived from the rat vasopressin gene sequence, were used to amplify lobster DNA to obtain suitable probes for screening a cDNA and genomic DNA library for the putative MIH gene sequence.

PCR amplification of both controls, i.e. rat genomic DNA and the rat vasopressin gene clone using an annealing temperature of 55°C and 1.5 mM $MgCl_2$, produced a single band, 1.2 kb long (Fig. 2.4, lanes 18-20), which when cleaved with restriction enzyme *Bgl*II, gave the expected two bands, 0.69 and 0.58 kb in length (lane 17). No visible amplification product of lobster genomic DNA was detected using similar conditions (lane 12), suggesting that optimal conditions had to be determined for the lobster when using the rat vasopressin primers.

PCR amplification of lobster genomic DNA at different $MgCl_2$ concentrations and at annealing temperatures of 45°, 50° and 55°C, produced a whole series of bands ranging from 0.95 to 0.4 kb in most samples (Fig. 2.4, lanes 3-16). Two predominant bands, 947 bp and 564 bp long, were seen in most samples; these bands were subsequently named the 947 PCR and 564 PCR bands as they ran at the same

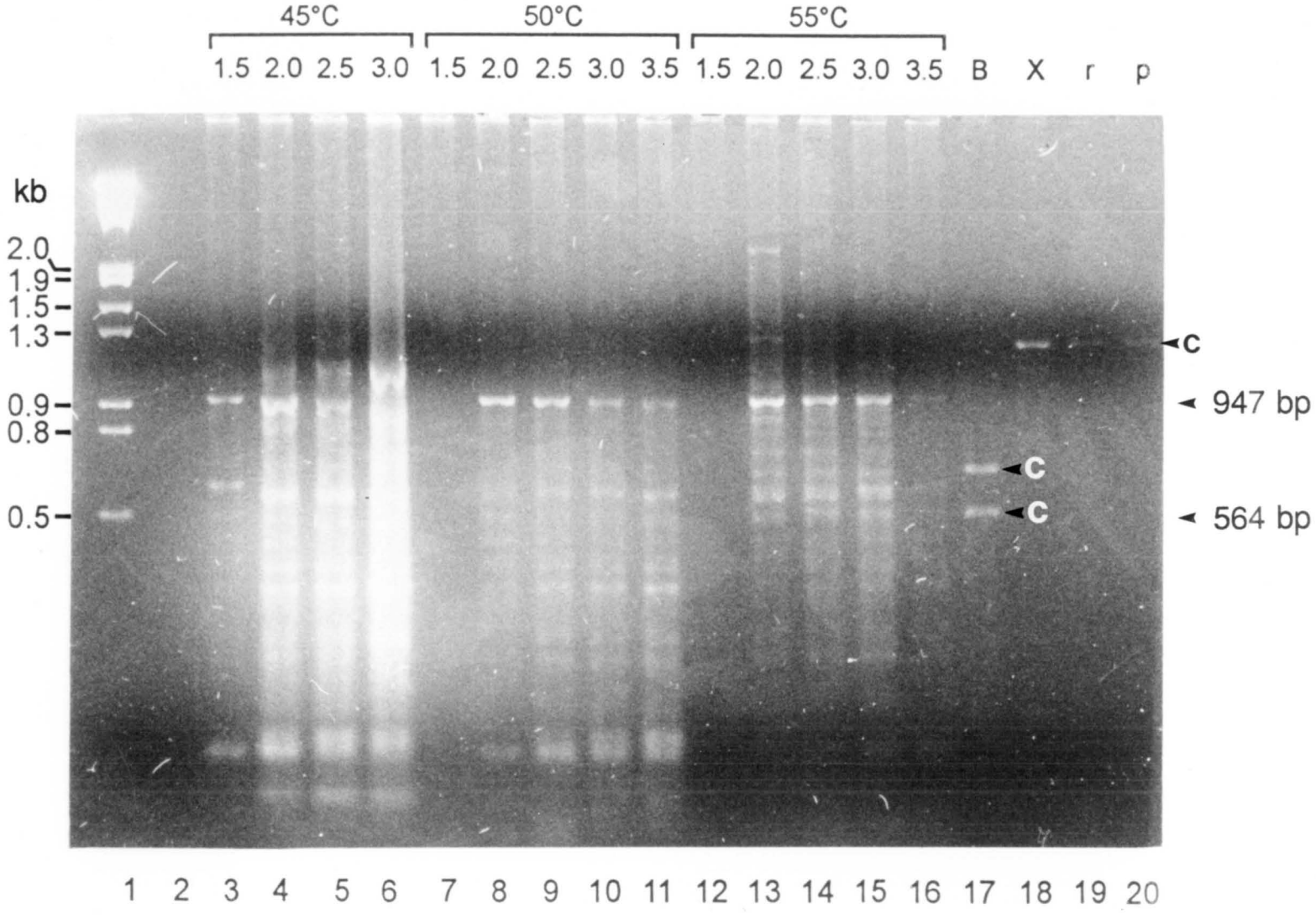
mobility as the 947 bp and 564 bp bands of the *Eco*RI and *Hin*dIII λ DNA molecular size markers. Several larger PCR products (2.1, 1.3, and 1.1 kb) were seen in some of the samples (Fig. 2.4, lanes 4-6, and 13).

Southern blot analysis of these PCR products using the rat vasopressin DNA as a probe revealed consistently the lack of hybridization signals to the 947 bp fragment (Fig. 2.5, Table 2.1).

Table 2.1. Molecular weights of PCR products detected by ethidium bromide staining (UV visible) in the gel (Fig. 2.4), and PCR products detected by Southern hybridization to the rat vasopressin probe (hybridization positive bands) (Fig. 2.5), The major PCR products, as detected by ethidium bromide, the 947 bp and 564 bp bands (0.95 and 0.56 kb) are indicated in italics. The sizes of the bands were determined by comparison to the *Eco*RI and *Hin*dIII digested λ DNA molecular size markers.

Annealing temperature	mM $MgCl_2$	UV visible bands (kb)	Hybridization positive bands (kb)
45°C	1.5 mM	0.98, 0.66	1.26
	2.0 mM	1.1, *0.95*, 0.61, *0.56* 0.5, 0.44, 0.40	1.21, 0.90, 0.64 0.51
	2.0 mM 3.0 mM	1.1, *0.95*, 0.61, *0.56*, 0.44, 0.40	absent
50°C	1.5 mM	absent	absent
	2.0 mM	*0.95*, 0.61, *0.56*, 0.44, 0.40	1.26, 1.07, 0.90
	2.5 mM	*0.95*, 0.61, *0.56*, 0.44, 0.40	0.90
	3.0 mM 3.5 mM	*0.95*, 0.61, *0.56*, 0.44, 0.40	absent
55°C	1.5 mM	absent	absent
	2.0 mM	2.1, 1.3, *0.95*, 0.61, *0.56*	(1.26, 1.07, 0.90) seen as 1 band, 0.5
	2.5 mM	*0.95*, 0.61, *0.56*	1.26, 1.07, 0.90, 0.5
	3.0 mM	*0.95*, 0.61, *0.56*	1.26, 1.07, 0.90, 0.5
	3.5 mM	*0.95*, 0.61, *0.56*	absent

Figure 2.4. PCR amplification of lobster DNA. The varying conditions of annealing temperatures and $MgCl_2$ concentrations are indicated above the gel. Ten µl of each 50 µl lobster PCR reaction was loaded on the gel (lanes 3-16). Lanes 17 and 18 contained 1 µl of rat PCR product (25 µl) digested with *Bgl*II (B), and *Xmn*I (X). Lanes 19 and 20 contained 0.5 µl of rat (r) and 0.05 µl of plasmid pBS 13+ (p) PCR products respectively. The positions of the 947 bp and 564 bp PCR products are indicated by ◂, and bands from the PCR controls are indicated (C). *Eco*RI and *Hin*dIII digested λ DNA size markers are indicated on the left (kb): 2.0, 2,027 bp; 1.9, 1,904 bp; 1.5, 1,584 bp; 1.3, 1,375 bp; 0.9, 947 bp; 0.8, 831 bp; 0.5, 564 bp.


45°C
50°C
55°C
1.5 2.0 2.5 3.0 1.5 2.0 2.5 3.0 3.5 1.5 2.0 2.5 3.0 3.5 B X r p
kb
2.0
1.9
1.5
1.3
0.9
0.8
0.5
c
947 bp
c
c
564 bp
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

However, the probe hybridized to bands (1.26, 1.07, 0.90, and 0.6-0.4 kb) which were not detected visually in the ethidium bromide stained gels, but the hybridization signal was very strong for PCR products at an annealing temperature of 55°C (Fig. 2.5, lanes 13-15). This suggests that the amplified DNA sequences seen in the ethidium bromide stained gel in Figure 2.4, had low homology to the rat vasopressin gene.

2.3.3 Re-amplification of PCR products

The major PCR products amplified using primers based on the rat vasopressin gene were isolated and re-amplified to obtain sufficiently pure PCR products which could be used as probes to isolate the putative MIH sequence from the lobster. Since the 0.95 kb and 0.56 kb fragments were detected most consistently, these fragments were gel purified and re-amplified using similar PCR conditions (Fig. 2.6). Re-amplification of DNA isolated from the 947 bp band produced a 947 bp band and a 0.56 kb band (lanes 5 and 7, Fig. 2.6). Amplification of DNA isolated from the 564 bp band sometimes produced two bands around 0.56 kb (lanes 4 and 6, Fig. 2.6). These results suggested that the initial 0.95 kb and 0.56 kb PCR products had either additional PCR primer binding sites within the sequence or, the DNA isolated from the bands also contained minute quantities of other lobster DNA sequences besides the major 947 bp and 564 bp DNA bands. Both the 947 bp and 564 bp bands were gel purified before use in subsequent experiments. The other PCR products which were detected by hybridization to the rat vasopressin gene were present at too low levels to be efficiently isolated and amplified.

2.3.4 DNA dot blot and Southern blot analysis of lobster and rat genomic DNA using the 564 bp and 947 bp PCR products as probes

To demonstrate that the PCR fragments were derived from lobster genomic DNA, dot blot and Southern blot analysis was carried out using the isolated PCR fragments as probes. The 947 bp PCR fragment hybridized very strongly to lobster genomic DNA (100 ng), and weakly to the same quantity of rat genomic DNA, in the dot blot analysis, using medium stringency conditions (Fig. 2.7).

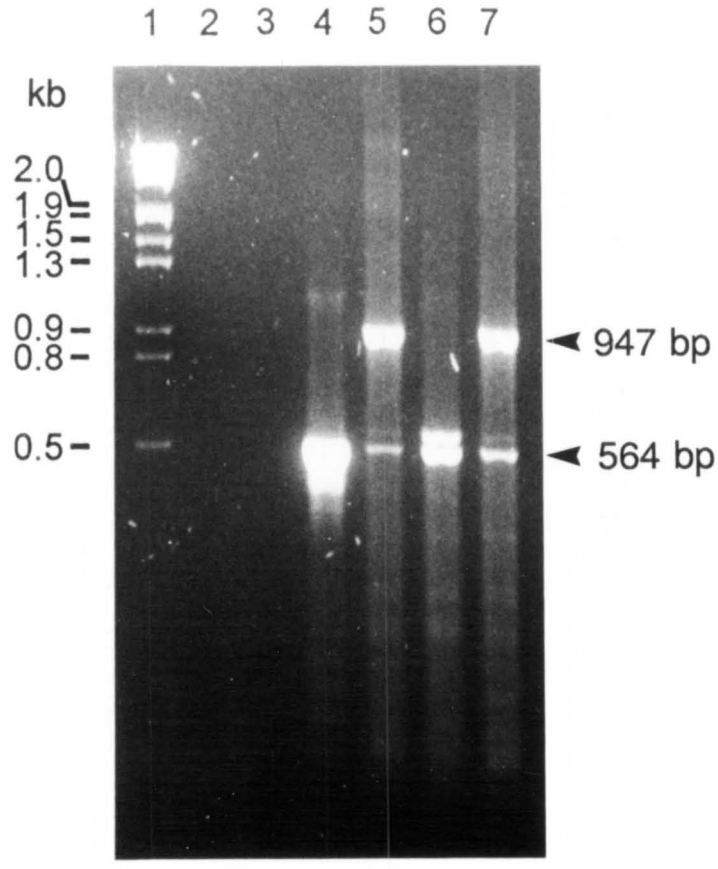


Figure 2.6. PCR re-amplification of DNA eluted from the 564 bp PCR band (lanes 4 and 6), and 947 bp band (lanes 5 and 7). Re-amplification of the 564 bp PCR fragment sometimes produced two bands (lane 6), whereas re-amplification of the 947 bp PCR fragment produced two bands at 947 bp and 0.56 kb (lanes 5 and 7). *Eco*RI and *Hin*dIII digested λ DNA size markers are indicated (kb).

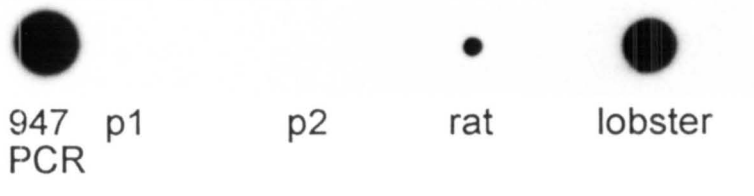


Figure 2.7. Genomic DNA dot blots. 100 ng of rat and lobster DNA, and 8 ng of the 947 bp PCR product were probed with 25 ng of the 947 bp PCR product. Also included were 200 ng of plasmid vector DNA: p1, pBS m13+; p2, pSPORT 1. The dot blot was washed in 0.5 x SSC/ 0.1% SDS at 65°C for 15 min, and exposed for autoradiography for 45 h.

When both PCR products were used as probes in Southern analysis of lobster and rat genomic DNA, the 564 bp fragment hybridized to a 4.1 kb band in a *Bgl*II digest of lobster DNA (Fig. 2.8a, lane 1) whereas the 947 bp fragment hybridized to three bands, 6.5 kb, 4.8 kb, and 3.6 kb (Fig. 2.8b, lane 1) in the same digest of that particular membrane. These results differ from the hybridization pattern shown in the Southern blot analysis of the same membrane when the rat vasopressin gene was used as a probe (Fig. 2.2a) where the rat vasopressin probe hybridized to two bands, 4.5 and 2.3 kb, in the *EcoR*I digest of lobster genomic DNA.

Although DNA dot blot analysis with the 947 bp fragment showed hybridization for both lobster and rat genomic DNA (Fig. 2.7), both the 947 bp and 564 bp fragments did not hybridize to restricted digests of rat genomic DNA using high stringency conditions (Fig. 2.8a and 2.8b, lanes 5-8).

The two PCR products also hybridized to undigested lobster DNA, observed as high molecular weight DNA near the gel wells (Fig. 2.8a and 2.8b, lanes 1-4). Several different restriction endonucleases were trialed; the restriction enzymes *Bgl*II and *Eco*RI gave better digestion. Although excess enzyme was used (12 U/μg DNA), and the period of incubation extended to 18 h, the lobster DNA was still incompletely digested. This could be due to polysaccharide contamination in the extraction of DNA from lobster gill tissue. The polysaccharide contaminant could have been removed by digestion of purified DNA with α-amylase, or avoided by extracting DNA from another type of tissue such as the hepatopancreas, or using the CTAB (Cetyltrimethylammonium bromide) method of DNA extraction commonly used for plant tissue (Towner, 1993).

2.3.5 *In situ* hybridization of lobster eyestalk

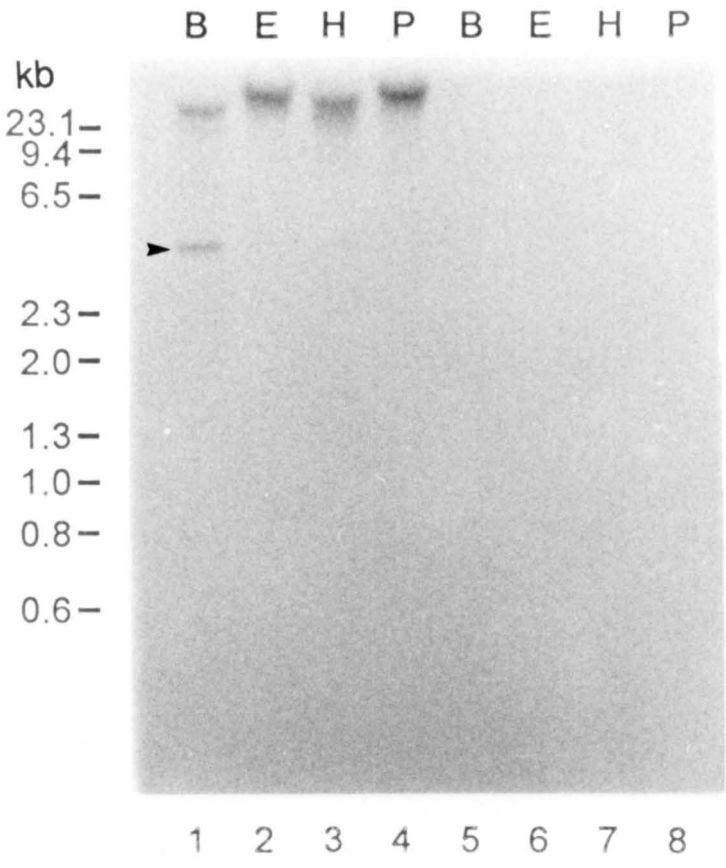
To determine whether the PCR products contained gene sequences which were expressed in the eyestalk, the site of production of the crustacean MIH, both the 564 bp and 947 bp PCR products were used as probes for *in situ* hybridization to lobster eyestalk sections. Both PCR products were found to hybridize to specific regions

Figure 2.8. Southern hybridization of lobster and rat genomic DNA to the 564 bp and 947 bp PCR products.

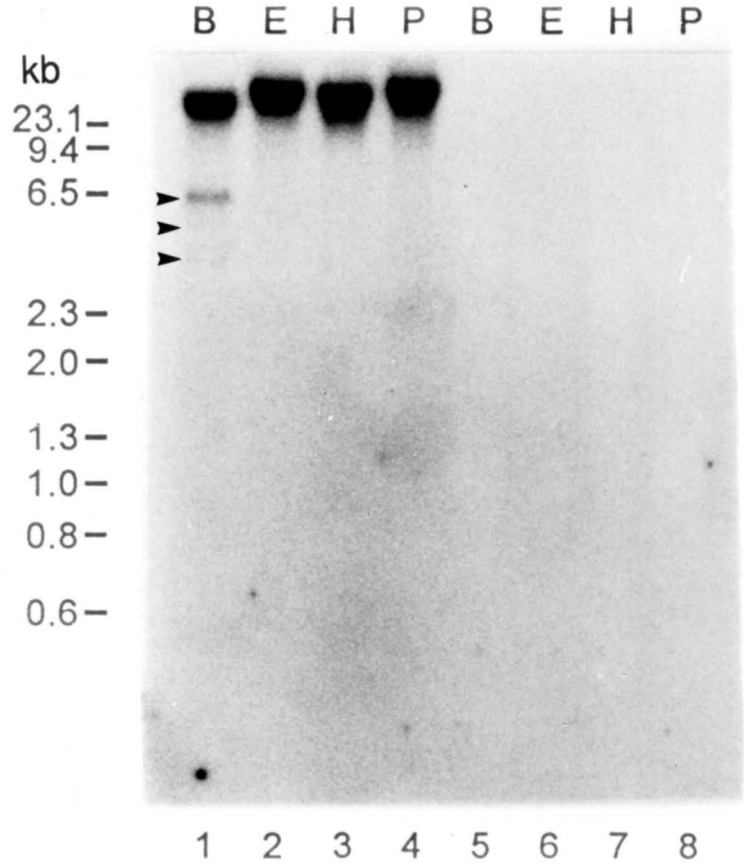
(**a**) The Southern blot of lobster and rat genomic DNA digested with *Bgl*II (B), *Eco*RI (E), *Hin*dIII (H), and *Pst*I (P) was stripped after probing with the rat vasopressin gene (Fig. 2.2), and hybridized to the 564 bp PCR product. The membrane was washed at high stringency (0.1 x SSC/0.1% SDS at 65°C for 15 min), and exposed for autoradiography for 10 days.

(**b**) Southern hybridization of lobster and rat genomic DNA to the 947 bp PCR product. The Southern blots shown in panel **a** was stripped and hybridized to the 947 bp PCR probe. After autoradiography for 3 days, the membrane was washed at a higher stringency (0.1 x SSC/0.1% SDS at 68°C for 30 min), and exposed for autoradiography for a further 10 days. The location of the DNA size markers (λ Drigest III) are indicated (kb): 23.1, 23,130 bp; 9.4, 9,416 bp; 6.5, 6,557 bp; 2.3, 2,332 bp; 2.0, 2,027 bp; 1.3, 1,353 bp; 1.0, 1,078 bp; 0.8, 872 bp; 0.6, 603 bp. Arrows denote the positions of the hybridization bands.

a


B E H P B E H P
kb
23.1
9.4
6.5
2.3
2.0
1.3
1.0
0.8
0.6
1 2 3 4 5 6 7 8

b


B E H P B E H P
kb
23.1
9.4
6.5
2.3
2.0
1.3
1.0
0.8
0.6
1 2 3 4 5 6 7 8

within the intermoult lobster eyestalk: neurosecretory cells of the medulla terminalis, medulla interna, medulla externa, and lamina ganglionaris ($p < 0.0001$, Table 2.2). The 947 bp PCR product was localized over the nuclei of densely packed, small neurosecretory cells (X-organs) above the medulla externa, at the junction between the medulla interna and terminalis, and the neurosecretory cells of the medulla terminalis (MTXO) and medulla terminalis ganglionaris X-organs (MTGXO) (Fig. 2.9a-e). Probe localization was also seen in the nuclei of some of the cells above the lamina ganglionaris, and over the nuclei of the larger neurosecretory cells adjacent to the X-organs of the medulla terminalis (Fig. 2.9d, f).

The 564 bp PCR product hybridized to the nuclei of the neurosecretory cells in the medulla terminalis ganglionaris (MTGXO) and the larger adjacent neurosecretory cells, and in the perikaryon of the neurosecretory cells of the medulla externa and medulla terminalis (data not shown). The accumulation of silver grains, however, was not as heavy as seen for the 947 bp PCR product, although both experiments were carried out simultaneously (Table 2.2).

2.3.6 Expression of the PCR encoded gene sequences in the premoult and intermoult stage

To determine whether the expression of the PCR encoded gene sequences were related to the moult cycle, *in situ* hybridization of the PCR products to eyestalk sections prepared from both an intermoult lobster, and premoult lobster were compared. The fixation and hybridization of eyestalk sections from an intermoult and premoult lobster were carried out simultaneously; eyestalk sections from the premoult and intermoult lobster were placed on either half of the same glass slide. Following exposure of the slides to liquid emulsion, the slides were developed simultaneously.

When the 947 bp and 564 bp fragents were hybridized to eyestalk sections taken from a lobster which was in the process of moulting (premoult - presence of a new exoskeleton underneath the outer cuticle), significantly lower densities of silver grains were seen over the neurosecretory regions of the eyestalk (Fig. 2.9f-i) (paired t-test; $p < 0.0001$ for 947 bp, $p < 0.005$ for 564 bp probe).

Figure 2.9. *In situ* hybridization of intermoult lobster eyestalk to the 947 bp PCR probe. (**a**) Whole mount of LS eyestalk showing probe hybridization to the neurosecretory cells of the lamina ganglionaris (LG), medulla externa (ME), medulla interna (MI) and medulla terminalis (MT). Probe localization was seen over the nucleus and cytoplasm of the neurosecretory cells above the medulla externa (**b**), and at the medulla externa and interna junction (**c**). Bars: a = 200 μm; b and c = 100 μm.

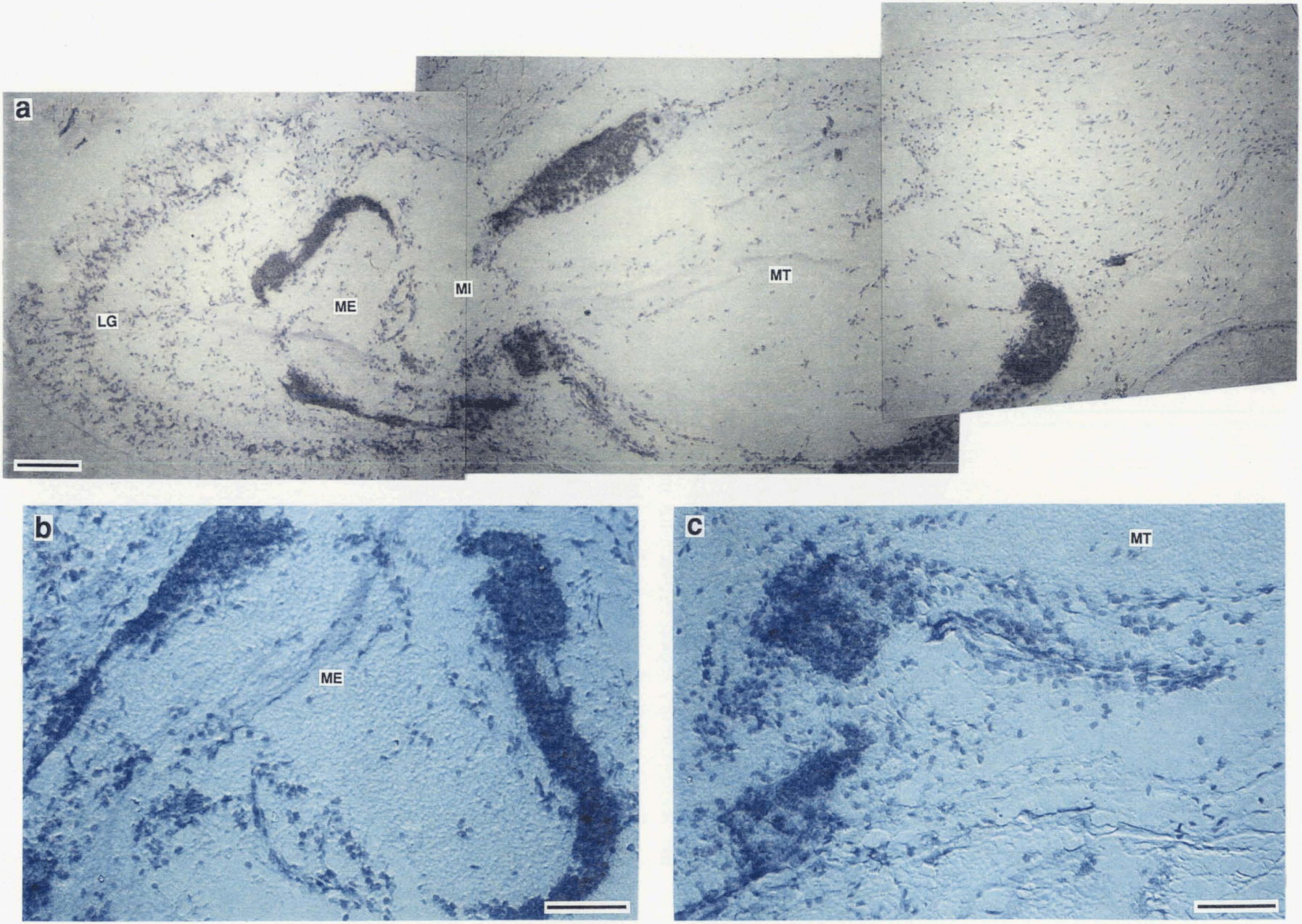
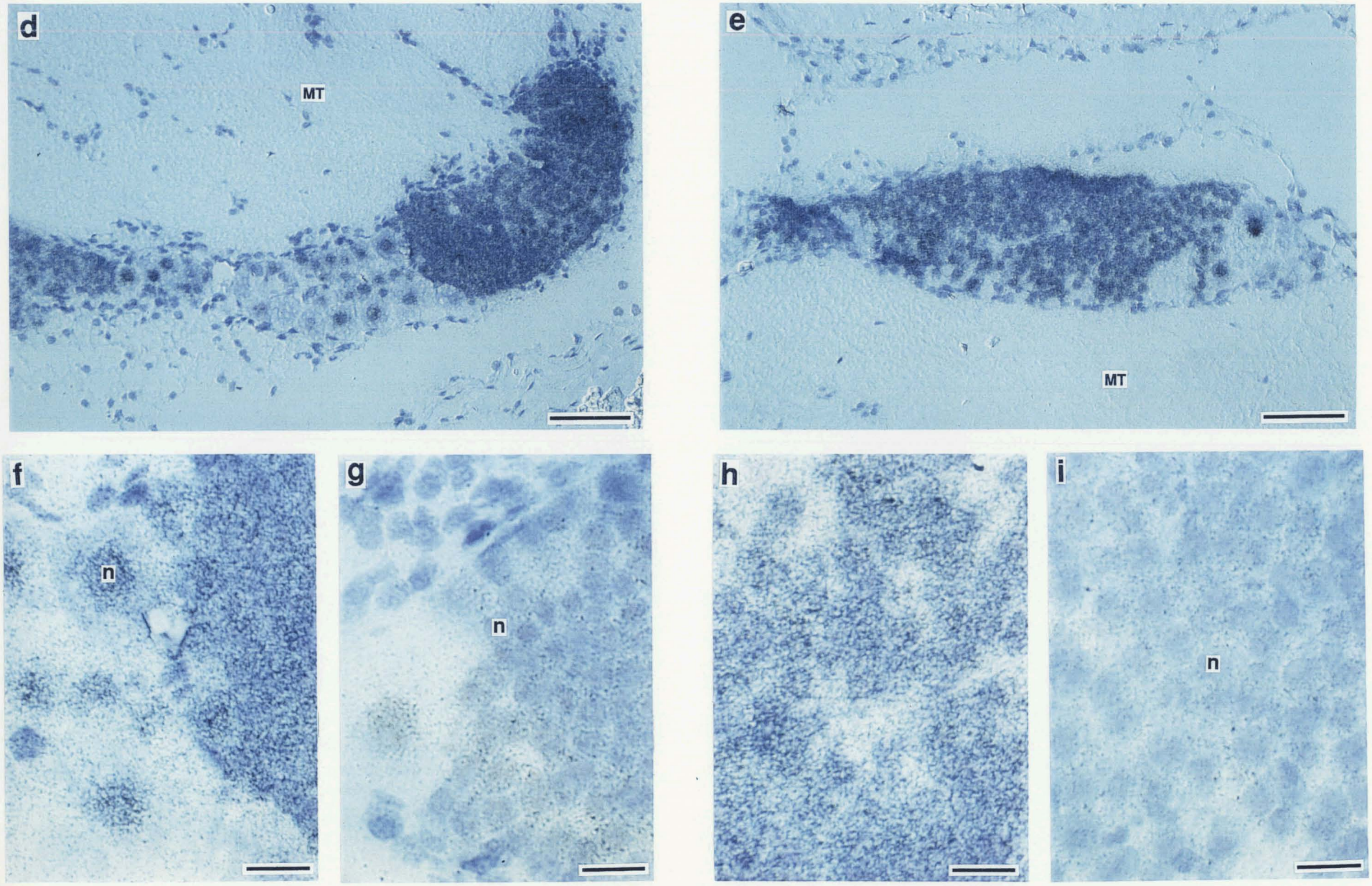

a
LG
ME
MI
MT
b
ME
c
MT

Figure 2.9. *In situ* hybridization of intermoult lobster eyestalk to the 947 bp PCR probe (cont.). The 947 bp PCR product hybridized to the nucleus and cytoplasm of the smaller, densely packed cells of the medulla terminalis ganglionaris X-organs, and the nucleus of the larger neurosecretory cells of the medulla terminalis (**d**, **f**), and the neurosecretory region between the medulla terminalis and medulla interna, possibly the sinus gland (**e**, **h**). Probe hybridization was significantly reduced in eyestalk sections taken from a premoult lobster, in the corresponding regions of the medulla terminalis (**g**), and sinus gland (**i**). The slide was exposed to emulsion for 4 days. Bars: d and e = 100 μm; f, g, h and i = 200 μm.


d
MT
e
MT
f
n
g
n
h
i
n

Estimates of the average number of silver grains over similar positions within the neurosecretory regions of the intermoult and premoult lobster eyestalk section are summarised in Table 2.2. In the premoult eyestalk, hybridization of the 564 bp probe was confined to the cytoplasm of the neurosecretory cells of the medulla terminalis, medulla externa, and the sinus gland, compared to both cytoplasmic and nuclear localization of the probe for the intermoult lobster eyestalk (data not shown). In the premoult lobster eyestalk section, the pattern of hybridization of the 947 bp probe was similar to the intermoult eyestalk section; however, probe hybridization was significantly reduced (Fig. 2.9f-i).

Table 2.2. Comparison of the average number of silver grains in a 64 μm^2 quadrat, taken from the neurosecretory regions of the medulla terminalis, medulla externa, and tissue designated as sinus gland (see Fig. 2.9). Silver grains were counted at a 1000 x magnification, in 10 random quadrats within 10 x 10 eyepiece micrometer grid, placed over the neurosecretory region of the eyestalk. The background level of silver grains was counted from tissue adjacent to the neurosecretory region.

Probe	Tissue	intermoult eyestalk		premoult eyestalk		paired t-test*
	region	n.s region	background	n.s region	background	
947 bp	medulla terminalis	96.6 ± 22.7	10 ± 4.5	65.6 ± 12.6	4.9 ± 4.3	p< 0.0001
	medulla externa	89.8 ± 17.6	16.7 ± 7.6	38.6 ± 13	3.2 ± 3.5	p< 0.0001
	sinus gland	104.7 ± 12.5	22.2 ± 6.8	44.5 ± 8.3	2.8 ± 1.2	p< 0.0001
564 bp	medulla terminalis	92.7 ± 15.8	5.2 ± 2.2	54.9 ± 13.1	3.5 ± 1.8	p< 0.001
	medulla externa	77.8 ± 12.1	13.3 ± 7.8	46.5 ± 17.0	4.8 ± 1.2	p< 0.005
	sinus gland	76.4 ± 13.5	6.8 ± 3.3	43.8 ± 10.8	0.9 ± 0.9	p< 0.001

* Paired t-tests comparison of the number of silver grains in the neurosecretory regions (n.s.) of the intermoult and premoult lobster eyestalk sections.

2.3.7 Northern blot analysis using the 947 bp PCR product as a probe

In situ hybridization of eyestalk sections demonstrated that the 947 bp fragment contained gene sequences which were expressed in the eyestalk. To determine

whether the gene sequence coded for a protein, i.e., by the ability of the probe to detect mRNA, and whether the mRNA was produced in different tissues, i.e., tissue specific expression, the 947 bp PCR product was used as a probe in northern blot analysis of mRNA and total RNA isolated from several different tissues of the lobster.

Total RNA isolated from the epithelia, gill, heart, hepatopancreas, and muscle of an individual lobster, and from the eyestalks of several lobsters was fractionated on a 1.2% agarose gel (Fig. 2.10a). Total RNA (25.8 ± 1.3 µg) isolated from the various tissues ran as a smear. At least 15 major RNA bands were visible in all the lanes (Fig. 2.10a, lanes 8-12). The 18S and 28S ribosomal RNAs (rRNA) were seen at 2.2 kb and 3 kb positions, respectively.

In contrast to total RNA, poly(A)$^+$ RNAs isolated from the different tissues were very distinct; poly(A)$^+$ RNA (1.82 ± 0.67 µg) ran as a smear along the lanes, with at least one distinct band in the epithelia, gill and heart. At least three distinct bands were observed in eyestalk tissue, and six bands in the hepatopancreas and muscle extract (Fig. 2.10a, lanes 1-6).

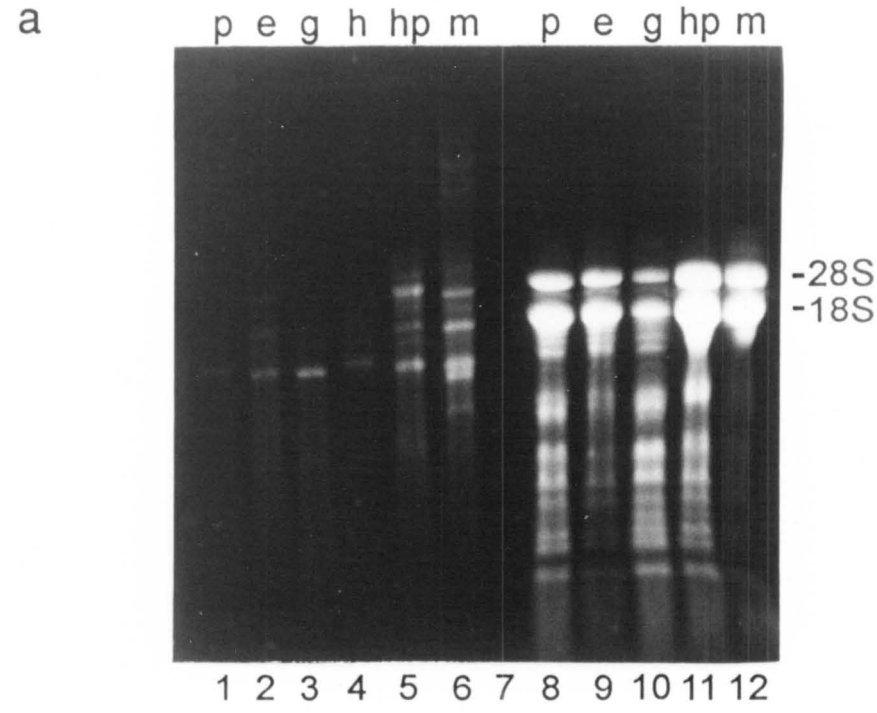
Northern blotting followed by hybridization using the 947 bp PCR fragment as a probe detected two bands (1.68 and 0.68 kb) in the poly(A)$^+$ RNA purified from the eyestalk (Fig. 2.10b, lane 2). The 947 PCR product also hybridized to two bands (1.8 and 0.3 kb) in the poly(A)$^+$ RNA from the heart, and three bands from the hepatopancreas (6.0, 2.2 and 0.90 kb), and muscle (5.3, 1.75 and 0.4 kb) (Fig. 2.10b, lanes 4-6). The presence of all these different sized mRNA bands suggests that the 947 PCR product could be detecting different mRNAs. The intensity of the signal was darker in the eyestalk compared to the other tissues, despite the fact that the amount of poly(A)$^+$ RNA loaded varied (eyestalk (1.44 µg)< muscle (2.43 µg) < hepatopancreas (2.7 µg)).

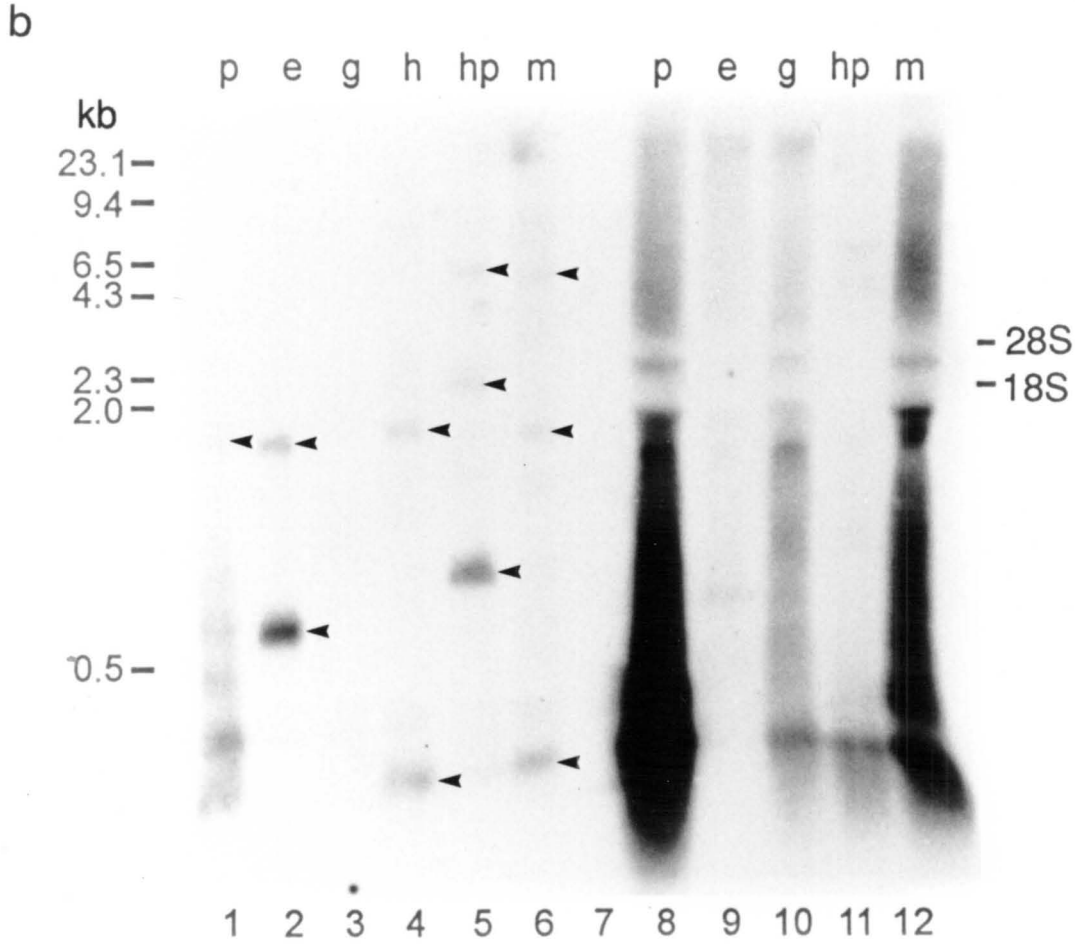
Hybridization bands on the autoradiogram were analysed semi-quantitatively by measuring the optical densitometry reading of each band with a scanner (AAB,

Figure 2.10. Northern blot analysis of poly(A)$^+$ RNA and total RNA isolated from several tissue types of the lobster.

(a). Electrophoretic analysis of poly(A)$^+$ and total RNA isolated from the lobster. Poly(A)$^+$ RNA (lanes 1-6) and total RNA (lanes 8-12) was isolated from the following tissue: *p*, epithelia; *e*, eyestalk; *h*, heart; *hp*, hepatopancreas; and *m*, muscle. The location of the 28S and 18S rRNA bands are indicated on the right.

(b). Northern hybridization of lobster poly(A)$^+$ and total RNA shown in panel a, to the 947 PCR probe. Poly(A)$^+$ (lanes 1-6) and total RNA (lanes 8-12) was isolated from the following tissue: *p*, epithelial; *e*, eyestalk; *h*, heart; *hp*, hepatopancreas; and *m*, muscle. The northern blot was washed in 0.2 x SSC/0.1% SDS at 60°C for 3 x 15 min, and exposed for autoradiography for 7 days at −80°C, with intensifying screens. ◂, indicate the locations of the bands. The location of single stranded *Hin*d III digested λ DNA size marker are shown on the left (kb): 23.1, 23,130 bp; 9.4, 9,416 bp; 6.5, 6,557 bp; 4.3, 4,361 bp; 2.3, 2,322 bp; 2.0, 2,027 bp; 0.5, 564 bp. The 28S and 18S rRNA bands are indicated on the right.


a
p e g h hp m p e g hp m
-28S
-18S
1 2 3 4 5 6 7 8 9 10 11 12


b
p e g h hp m p e g hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
-28S
-18S
1 2 3 4 5 6 7 8 9 10 11 12

Advanced American Biotechnology). The autoradiogram was scanned using three different sensitivity settings on the scanner. The readings were normalized to 1 µg quantities of poly(A)$^+$ RNA, and expressed as a percentage of the 0.68 kb eyestalk band for that particular sensitivity (Table 2.3). The relative intensities of the darkest bands in the other tissues as compared to the 0.68 kb eyestalk band were: heart, 80%; muscle, 43%; hepatopancreas, 41%; and the 1.68 kb eyestalk band, 33% of the 0.68 kb eyestalk band (Table 2.3).

Table 2.3. Relative intensities of northern hybridization bands (Fig. 2.10b). The autoradiogram was scanned using three different sensitivity settings. The optical densitometry measurement of each band was normalized to 1 µg, and expressed as a percentage of the 0.68 kb eyestalk band (n=3).

Tissue	Band size (kb)	Mean % ± s.d.
eyestalk	1.68	32.7 ± 6.8
	0.68	100
heart	1.80	24.5 ± 9.7
	0.30	76.9 ± 7.3
hepatopancreas	6.0	12.6 ± 6.0
	2.2	14.5 ± 7.1
	0.9	41.1 ± 3.6
muscle	5.3	16.3 ± 7.2
	1.75	12.6 ± 6.0
	0.40	43.0 ± 7.4

Northern blot analysis of total RNA isolated from epithelia, gill and muscle tissues showed non-specific hybridization; the final post-hybridization wash conditions (0.2 x SSC/0.1% SDS at 60°C) may have been too mild and hence, the probe remained bound to non-specific sequences in the total RNA. The high concentration of the 18S and 28S rRNA bands interfered with hybridization, as suggested by the absence of any hybridization signal at the position of these bands for several tissues. The sizes of the RNA bands were calculated by comparison to the

sizes of single stranded λ DNA markers run on the gel; hence the sizes (kb) given were only estimates.

Denatured RNA and DNA molecules of the same size migrate with similar mobilities and hence it was possible to use DNA restriction fragments as size markers (Williams and Mason, 1985). However, RNA migrated faster than DNA of equivalent sizes through agarose gels containing formaldehyde (Wicks, 1986). For northern blot analysis, formaldehyde was only included in the sample buffer.

2.3.8 Sequence analysis of the PCR products

2.3.8.1 Sequence analysis of the 947 bp PCR product

The nucleotide sequence and deduced amino acid sequence of the cloned 947 bp PCR product are shown in Fig. 2.11. By sequencing, the 947 bp PCR product was found to be 960 bp long, instead of the 947 bp as estimated by comparison to λ DNA molecular size markers in gel electrophoresis. The 947 PCR product showed several interesting features:

a) A stretch of repetitive sequence from nucleotides 249 to 440; the sequence AGTTGTAGTGTAGATGTATGCTGT which coded for amino acids Ser Cys Ser Val Asp Val Cys Cys was repeated eight times. A sequence similarity search to the nucleotide and protein sequences databases on the NCBI Blast E-Mail Server (National Center for Biotechnology Information) was done. The repetitive sequence, when translated in reading frame 3+, showed similarity to a whole range of metallothionein protein sequences (62% of the 100 homologous entries).

b) A possible splice junction, with the consensus sequence (GT-AG) (Mount, 1982) at possible exon-intron boundaries in reading frame 3+, (highlighted in Fig. 2.11). This corresponded to the protein coding region prediction, conducted on DNASIS, which suggested nucleotides 200-600 could include a protein coding region, whereas nucleotides 600-960 was non-coding (algorithm calculations were based on the theory developed by Fickett, 1982).

```
  1 AA CTG CCC AAG AGG AGG CAC ACC TGT ACA GTA AGT AGT ACA TCG CTT GTG
       Leu Pro Lys Arg Arg His Thr Cys Thr Val Ser Ser Thr Ser Leu Val

 51 CTC GTA CGC CAC CTG CCG GAC GCG GGA AGC GGC CTG CAG ATG GAG CAG
    Leu Val Arg His Leu Pro Asp Ala Gly Ser Gly Leu Gln Met Glu Gln
                                                      → intron
 99 TGT TCA CCC CTT GAG CAC GAC GGT ACA GCC CTG GGG GTATGATGACCTGGCCCT
    Cys Ser Pro Leu Glu His Asp Gly Thr Ala LEU Gly

153 TCACCCGACCCTTAAGGGTCCCCTTTTTTTTAATTGGGTGGTTAGTAAGGGGGGAAGAAGGGATAG
                              → exon
219 GATAGGGAAGGTAAAGTGGAG AAA CAA GAC*AGT TGT AGT GTA GAT GTA TGC TGT
                              Lys Gln Asp Ser Cys Ser Val Asp Val Cys Cys

273 AGT TGT AGT GTA GAT GTA TGC TGT AGT TGT AGT GTA GAT GTA TGC TGT
    Ser Cys Ser Val Asp Val Cys Cys Ser Cys Ser Val Asp Val Cys Cys

321 AGT TGT AGT GTA GAT GTA TGC TGT AGT TGT AGT GTA GAT GTA TGC TGT
    Ser Cys Ser Val Asp Val Cys Cys Ser Cys Ser Val Asp Val Cys Cys

369 AGT TGT AGT GTA GAT GTA TGC TGT AGT TGT AGT GTA GAT GTA TGC TGT
    Ser Cys Ser Val Asp Val Cys Cys Ser Cys Ser Val Asp Val Cys Cys

417 AGT TGT AGT GTA GAT GTA TGT TGT*AGA TCT TGT TGT CTT GAA GTG TCA
    Ser Cys Ser Val Asp Val Cys Cys Arg Ser Cys Cys Leu Glu Val Ser

465 ATG CTT TGC TGT ATA TGT ATT TTT GGG AAC TTT ATA TAC ACA GCC CGG
    Met Leu Cys Cys Ile Cys Ile Phe Gly Asn Phe Ile Tyr Thr Ala Arg

513 CAT ACA CAG CCT AGG GTG TAA TGCAGTGTTGCTAGGCTGGATTACAACAGCCTCTGT
    His Thr Gln Pro Arg Val stop

570 ATATATACGCTAGGCTAAAAGAGGTGTTCGTTCATTGTGTATATCGATTGATGACGTGCTTCTC

634 TCTGAGGCTCCGTGGGAACCCAATGTGTTCCATAGTTTGTGCTGGTGTCCTGATCTTGGATTGC

698 TGTGCTGGCCTGTGTCTGCAGGAGGGAGGCTAGGCTGGCTGCAGAGCCAGAGAGCCGCAGCCAT

767 AGCGGCACTGGTGGCTCATCCAGCCATCAGAAGCCGCTCGACGCTACATGTAAACAAACGCTGC

826 CCTCTGGAAGCAGCGCCCCCATTGGTTCCTTACGAGCCAACGTCTGAAAGCACCGCCCCCATTG

890 GTTCCTTAAACCGACTTCTGGAAGCACCGCCCCCATTGGTCCCGCTGTAGCCAGTGTCTCCCTG

954 CGGCCCT 3'
```

Figure 2.11. Nucleotide sequence and deduced amino acid sequence of the 947 PCR product. The positions of PCR primer V1, at the beginning of the sequence, and the complementary sequence of V2, at the 3′ end are underlined. The beginning and end of the DNA repeats is indicated by *, with the basic repeat unit underlined. Possible exon/intron boundaries are indicated by the GT-AG consensus sequences which are in **bold-face type.**

2.3.8.2 Sequence analysis of the 564 bp PCR product

By sequencing, the 564 bp PCR product was found to be 568 bp long. The nucleotide sequence and possible deduced amino acid sequence of the 564 PCR product is given in Figure 2.12. Prediction of the amino acid sequence was based on DNA sequence analysis for open reading frames, and protein coding region analysis performed on DNASIS, which predicts that nucleotides from 240 to 568 form a protein coding region.

The 0.56 kb PCR product obtained from the 947 PCR fragment re-amplification was probably caused by contamination of trace amounts of lobster DNA running at the same position as the 947 PCR fragment. Partial sequence analysis of this 0.56 kb band showed that it is closely related to the 564 PCR product (94% sequence homology).

2.3.8.3 Sequence homology search

Homology searches for both the 947 PCR and 564 PCR sequences were conducted to determine whether the PCR sequences were related to published MIH sequences from other crustacean species, and also how related these PCR products were to the rat vasopressin gene sequence.

Homology searches conducted using the NCBI Blast and DNASIS programs (Table 2.4) show that the 947 PCR product shared 42.6% sequence identity over a 954 bp overlap, with the corresponding sequence on the rat vasopressin gene, and the 564 PCR product shared 44.5% sequence identity over a 573 bp overlap with the rat vasopressin gene. This was not unexpected as this region within the rat vasopressin gene contained mostly intron sequence.

Comparison of the nucleotide sequences of the 947 PCR and 564 PCR product to published cDNA sequences of other crustacean eyestalk neuropeptides revealed the following results: the 947 PCR product shared 46-48% homology, over a >338 bp overlap, with the MIH cDNA sequences isolated from the shrimp *P. vannamei*, and shore crab, *C. maenas* respectively, whereas the complementary sequence of the

```
  1 AACTGCCCAAGAGGAGGCCAGTGGTTCGTCTTAGGGTCCTCTTCTCATTTTATCGTACGTCTTCT

 66 GTAGCAGGCGGGGCAAGGTTCTCCCTCCCGGGGGGATGTCAGACGTGTGGTGTTGGTGCTGGAGT

131 TTCTGTAAACCAGGATCCTGGGACGTTGAAAGCACCGACAAATAATATAAGCTGAATTTCAAACT

196 GCTTTTACGGTCTTTGTCTGTCACGTAAGGAAGCAGGAGCCAGCTCAGCG ATG CCC CCG TCG
                                                      Met Pro Pro Ser

258 GGC GTT CTT CTG GTT CCC TTG GTG TGT TCG TTG TTG TTT CTT GTG GTT
    Gly Val Leu Leu Val Pro Leu Val Cys Ser Leu Leu Phe Leu Val Val

306 CTC TTT CCC CTT TCG CCC TTT CTC CAT TTC TTC GTC CGT CAC TTG CTT
    Leu Phe Pro Leu Ser Pro Phe Leu His Phe Phe Val Arg His Leu Leu

354 TCC CTG TTG TTT CTT ACA TTG CAG CTG TTT CTT GTG TTT CTC TCC CTT
    Ser Leu Leu Phe Leu Thr Leu Gln Leu Phe Leu Val Phe Leu Ser Leu

402 GGT CAG CCG TGG CTC TTT CTT CTC CTC TTC TTT GGC CAG CTG CCG TGG
    Gly Gln Pro Trp Leu Phe Leu Leu Leu Phe Phe Gly Gln Leu Pro Trp

450 CTC CTC TTC CTG CTC GGC CAA CAG CTG CTG CTG CTG TTC TTT CAG TGC
    Leu Leu Phe Leu Leu Gly Gln Gln Leu Leu Leu Leu Phe Phe Gln Cys

498 AAT CTT CCT TAT GAA GTA GCT CTC AGG ATA CAG TTC CCT GTG TTT CTC
    Asn Leu Pro Tyr Glu Val Ala Leu Arg Ile Gln Phe Pro Val Phe Leu

546 GTC TTC TGT CTC CCT GCG GCC CT 3'
    Val Phe Cys Leu Pro Ala Ala
```

Figure 2.12. Nucleotide sequence and deduced amino acid sequence of the 564 PCR product (568 bp long). The positions of PCR primer V1, at the 5′ end of the sequence, and complementary sequence of V2, at the 3′ end are underlined.

947 PCR product shared 44% homology over a 917 bp overlap, to the CHH A cDNA sequence from the crayfish, *O. limosus*. The 564 PCR product also showed relatively high homology, 47% homology over a 446 bp overlap, to the shrimp MIH cDNA (Table 2.4).

Table 2.4. Sequence similarity of PCR products, 564 and 947, showing the % sequence identity shared along a stretch of nucleotide sequence, indicated as the bp overlap.

PCR product	Animal	Sequence	% sequence identity	bp overlap[a]
947 lobster	rat	vasopressin (480-1720)	42.6%	954
	lobster	564 PCR	44.8%	563
	C. maenas[b]	MIH cDNA	48.5%	338
	P. vannamei[c]	MIH cDNA	46.1%	414
	Orconectes limosus[d]	CHH A cDNA	44.1%	917
	mouse	metallothionein MT-1 gene	43.6%	906
	human	metallothionein 1-B gene, ex.1,2	43.4%	793
	sea urchin	metallothionein MT-A	44.0%	639
564 lobster	rat	vasopressin (480-1720)	44.5%	573
	P. vannamei[c]	MIH cDNA	47.3%	446
	rat	metallothionein MT-2, MT-1	44.3%	517

[a] The bp overlap refers to the length of a continual stretch of nucleotides within the sequence where homology was detected.
[b] cDNA sequence encoding both the signal peptide and MIH from *Carcinus maenas* (Klein *et al.*, 1993b).
[c] *Penaeus vannamei* (Sun, 1994).
[d] CHH cDNA precursor in *Orconectes limosus* (de Kleijn *et al.*, 1994a) to complementary sequence of the 947 PCR product.

When translated into a putative protein, the 947 PCR product also showed consistent similarity to the metallothionein amino acid sequence isolated from several vertebrate sources on searches conducted using the BLAST tool. The homologous protein region was encoded by the repetitive sequence of the 947 PCR product. In this region, the putative amino acid sequence showed 39% sequence identity (16/41 amino acids) to the mouse metallothionein amino acid sequence, and 34% sequence identity to the human metallothionein amino acid sequence. When compared to the nucleotide sequence of the metallothionein gene from the mouse and rat, both PCR products show approximately 44% sequence identity over a 517 bp overlap (564 bp sequence) and 906 bp overlap (947 bp sequence) respectively (Table 2.4).

2.4 DISCUSSION

The aim of this chapter was to isolate by PCR the putative MIH gene sequence from the lobster, *J. edwardsii*, using PCR primers derived from the conserved regions of the rat vasopressin gene. Mattson and Spaziani (1985a) provided both physiological and immunological evidence which suggests that the crustacean MIH is structurally related to the vasopressins.

The presence of vasopressin-like gene sequences in the lobster was initially confirmed by Southern hybridization to the rat vasopressin gene, using high stringency wash conditions. In the PCR of lobster DNA, using primers derived from the rat vasopressin gene, two PCR products were consistently amplified under various PCR conditions. These two PCR products, the 947 bp major band and 564 bp minor band, were selected and re-amplified for further analysis. Southern hybridization of lobster and rat genomic DNA using the 947 and 564 PCR fragments as probes, showed that these amplified sequences were derived from the lobster.

To determine whether these PCR products contained protein coding sequences which were somehow related to crustacean MIH, the 947 and 564 PCR fragments were used as probes for northern blot analysis, and *in situ* hybridization of lobster

eyestalk sections, i.e., the site of production of the crustacean MIH. These experiments showed that the 947 PCR product contained protein coding sequences. Northern blot analysis of poly(A)$^+$ RNA isolated from different tissues of the lobster gave hybridization bands in poly(A)$^+$ RNA isolated from the epithelia, eyestalk, heart, hepatopancreas, and muscle tissue, using medium stringency wash conditions, and a 7 day exposure. The signal was stronger for the eyestalk, where less poly(A)$^+$ RNA was analyzed in comparison with the hepatopancreas and muscle, suggesting that the 947 PCR product could contain sequences coding for the putative MIH neuropeptide in the eyestalk. Sequence analysis of the 947 PCR fragment revealed that the PCR product contained putative protein coding regions, and possible intron/exon boundaries which would suggest that this sequence was part of a structural gene.

The 947 PCR product could also be detecting other related mRNA transcripts in addition to the putative MIH sequence in the eyestalk. Detection of several mRNA transcripts of varying sizes in the different tissues by the one probe suggests that either the 947 PCR product could be detecting the products of tissue specific alternate splicing of a common precursor, or related sequences. Most nuclear mRNA precursors (pre-mRNAs) in higher eukaryotes contain multiple introns which must be precisely excised by RNA splicing. Some pre-mRNAs are alternatively spliced in different cell types or at different times during development. Hence, regulated alternative splicing can lead to the production of different proteins from a single pre-mRNA, or can function as an on-off switch during development (Amara, 1985; Maniatis, 1991; McKeown, 1992; López, 1995).

In situ hybridization of intermoult lobster eyestalk sections using the 947 bp and 564 bp PCR fragments as probes revealed that both PCR products contained sequences which were specifically expressed in the neurosecretory cells of the medulla externa, interna, and terminalis. Nuclear localization of the probes was also observed in the retinal cells, and above the lamina ganglionaris.

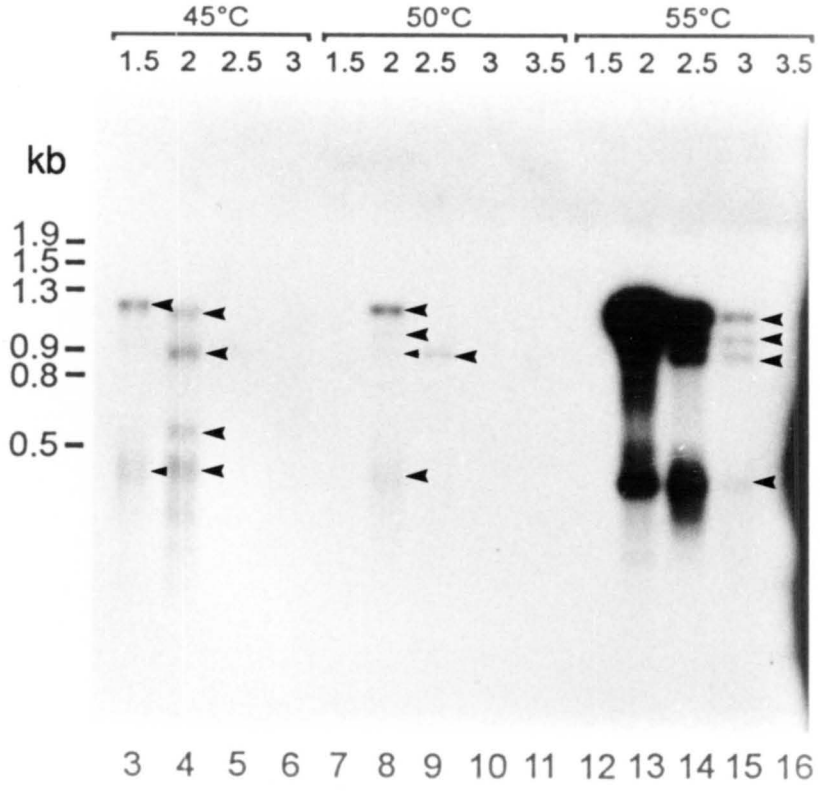


Figure 2.5. Southern blot analysis of the lobster PCR products shown in Fig. 2.4 (lanes 3-16), hybridized to the rat vasopressin probe, pBS 13+. The annealing temperatures and $MgCl_2$ concentrations are indicated. The final post-hybridization wash used was 0.5 x SSC/0.1% SDS at 65°C for 15 min, and the membrane was exposed to X-ray film for 6 days at -70°C. Lanes 17-20 from Fig. 2.4 were not included as they were highly overexposed. The location of DNA size markers (*Eco*RI and *Hin*dIII digested λ DNA) are indicated (kb). ◂, location of the bands hybridized to the rat vasopressin gene.

These results differed from the expression of putative MIH in the shore crab *C. maenas*, where putative MIH was detected at the protein and mRNA level, in a subset of cells of the medulla terminalis and in the sinus gland (protein), using both immunocytochemical and *in situ* hybridization techniques (Dircksen *et al.*, 1988; Klein *et al.* 1993a). However, by testing various sections of the eyestalk neurosecretory system for MIH-like activity, Webster (1986) demonstrated that in addition to the medulla terminalis, both the medulla externa and medulla interna had a small effect on the inhibition of ecdysteroid synthesis by *Carcinus* Y-organs.

Preliminary observations suggested that the expression of both the 947 PCR and 564 PCR fragments could be related to the moulting cycle; the hybridization of both PCR fragments were significantly reduced in eyestalk sections taken from a lobster in the premoult stage of the moulting cycle, compared to the intermoult lobster. Durand (1956) found that the neurosecretory activity of a certain cell type (type 2) in the X-organs of *O. limosus* changed in relation to the moulting cycle, which led him to conclude that this cell type produces MIH. Dircksen *et al.* (1988) also found that MIH immunopositive perikarya consistently exhibited faint immunostaining in the eyestalk of premoult *C. maenas* (stage D2). However, no striking moult cycle related changes in the immunostainability of the *C. maenas* sinus gland for MIH and CHH were observed. It is possible that no moult cycle changes were detected as synthesized proteins were released into the haemolymph, whereas mRNA remained in the cytoplasm. Therefore, there may not be any build-up of MIH in the eyestalk, before and after the moulting event.

Based on *in situ* localization of the 947 PCR fragment to the lobster eyestalk sections, and circumstantial evidence which suggested that the expression was somehow related to the moult cycle, the 947 PCR fragment seemed a likely candidate as a probe to search for the putative MIH gene sequence in the lobster, *J. edwardsii*. Sequence homology searches to determine how related the 947 PCR sequence was compared to published putative MIH sequences from other crustacean species showed ca. 46-48% homology to cDNA sequences encoding the putative MIH isolated from the shrimp, *P. vannamei*, and crab, *C. maenas*. The observed homology (44-46%)

could very well reflect interspecific differences; there was 49.4% nucleotide sequence homology over 324 bp overlap between the putative MIH isolated from *P. vannamei* and *C. maenas*, and 25% homology in the amino acid sequence of the putative MIH isolated from *C. maenas* and the lobster, *H. americanus*. The sequence homology could also be due to the fact that the 947 PCR fragment was derived from genomic DNA and therefore, contains intronic sequences. Comparison of the predicted protein coding region of the 947 PCR sequence with the putative MIH cDNA sequence from *C. maenas* and *P. vannamei* gave ca. 49% sequence identity over a 187 and 199 bp overlap, respectively.

Comparison of the 947 PCR sequence to the rat vasopressin gene showed 42.6% over a 954 bp overlap. This would account for the positive hybridization of the 947 PCR fragment to rat genomic DNA using medium wash stringency conditions in DNA dot blot analysis, and the absence of hybridization signals using high wash stringency conditions in Southern hybridization of digested rat genomic DNA.

Comparison of the putative protein sequence translated from the 947 PCR fragment, to amino acid sequences in the protein database, showed homology to several metallothionein sequences (62 of the 100 entries retrieved) isolated from both vertebrates and invertebrates sources. Closer examination of the 947 PCR sequence revealed that the homologous amino acid sequence was encoded by the repetitive sequence, from nucleotides 249-440. The putative amino acid sequence encoded by this region contained a high proportion of cysteine residues which had a fixed distribution along the amino acid sequence (24 out of 63 amino acid residues). The aromatic amino acids (Phe, Tyr, Trp) were also lacking in this region. These two features are characteristic of the metallothionein proteins (Haurowitz, 1963; Rainbow, 1988).

The metallothioneins are metal binding proteins which protect cells against excess concentrations of heavy metals. The gene is expressed at basal levels, but is induced to greater levels of expression by heavy metal ions (Bonwick *et al.*, 1991;

Lewin, 1994; Radtke *et al.*, 1995). Copper, bound in metallothioneins, is transferable to the respiratory pigment haemocyanin. In decapod crustaceans, haemocyanin levels vary during the moult cycle (Zuckerkandl, 1960; Djangmah and Grove, 1970; as cited in Rainbow, 1988). The copper appears to be stored as metallothionein complex in the hepatopancreas during the moult, and is subsequently released when new haemocyanin is synthesized (Engel *et al.*, 1985; Brouwer, 1986; cited in Rainbow, 1988). The observed differences in the expression of the 947 PCR sequence in the premoult and intermoult eyestalk were probably not related to the expression of the metallothionein gene, as in the blue crab *C. sapidus*, the estimated concentrations of metallothionein were highest during intermoult and premoult, and lowest during the papershell stage (Engel, 1987).

Although nucleotide sequence analysis of the 947 PCR product suggested it could be related to the metallothionein protein, it is more likely to be related to the MIH characterized by Mattson and Spaziani (1985a) based on its homology to the rat vasopressin gene, and expression in the eyestalk, the classical site of MIH production. Based on the evidence that the PCR products could detect genes expressed in the eyestalk, and the expression of these genes appeared to be related to the moult cycle, the 947 PCR fragment was selected for further use as a probe to isolate the putative MIH gene sequence from a cDNA library and a lobster genomic DNA library (Chapter III).

CHAPTER III

Cloning and sequence analysis of the putative MIH gene from the lobster *Jasus edwardsii*

3.1 INTRODUCTION

PCR was used to amplify the 947 PCR fragment which might contain sequences representing the putative MIH gene sequence from lobster DNA (Chapter II). Analysis of the PCR product suggested that this 947 PCR DNA fragment may be related to the MIH characterized by Mattson and Spaziani (1985a). The 947 PCR product contained protein coding sequences which were expressed in the lobster eyestalk, as demonstrated by northern blot analysis, and *in situ* hybridization using the 947 PCR product as a probe.

Sequence analysis of the 947 PCR sequence revealed over 40% sequence similarity to the rat vasopressin gene. The expression of this sequence in the eyestalk appeared to be related to the moulting cycle. Thus the 947 PCR product was selected as a probe to isolate the putative MIH gene sequence from both the lobster genomic DNA and eyestalk cDNA library.

The 947 PCR product was used in conjunction with the rat vasopressin gene to retrieve two clones, p2A and p66, with 7.4 and 8.0 kb inserts, respectively, from the lobster genomic library. Subsequently, the 947 PCR sequence was used in conjunction with the genomic clone, p66, to isolate three different cDNA clones, peJK1, peJK2 and peJK3 from the lobster eyestalk cDNA library. This chapter describes the isolation, the analysis of the three cloned cDNA sequences, and how the three compare to putative MIH cDNA sequences from other crustacean species.

3.2 MATERIALS AND METHODS

3.2.1 Isolation of the putative MIH gene sequence from the lobster genomic DNA library

(i) Construction of the lobster genomic DNA library

High molecular weight lobster DNA was partially digested with *Sau*3A, a frequent 4 base cutter, to produce a set of random overlapping sequences between 15-23 kb, as described in Maniatis *et al.* (1982). The digested DNA was fractionated on a 0.4% agarose gel in 1 x TAE (0.04 M Tris-acetate, 1 mM EDTA) and DNA between 15-23 kb was recovered by electroelution using the BIO-TRAP apparatus (Schleicher and Schule).

The *Sau*3A ends of the genomic inserts were partially filled in with dATP and dGTP and cloned into the *Xho*I half-site arms of lambda GEM-11, a λ cloning vector capable of accepting 9-23 kb inserts, according to the supplier's instructions (Promega). Based on trial ligations, a ratio of 1:2, vector:insert DNA was used in the ligation reactions. The *cos* sites of the vector were preannealed by incubation in 10 mM $MgCl_2$ for 1 h at 42°C (Kaiser and Murray, 1985), and subsequently λ DNA was packaged *in vitro* (Rosenberg, 1987).

Packaged phages were plated using freshly prepared plating cells of *E. coli* strain KW 251 (Appendix 1), on tryptone broth (TB) agar plates as described by Sambrook *et al.* (1989) and Silhavy *et al.* (1984). Eight small libraries (10^3-10^4 pfu/μg DNA) were constructed, and these were amplified and titred prior to screening, according to Sambrook *et al.* (1989).

(ii) Screening for the putative MIH gene sequence using the rat vasopressin gene and the 947 PCR sequence

The amplified genomic libraries were screened for the putative MIH gene sequence using the *in situ* plaque hybridization technique of Benton and Davis (1977) as described in Sambrook *et al.* (1989). Nylon filters were hybridized to 250 ng of [α-^{32}P] dCTP pBS 13+ (rat vasopressin gene) nick translated probes, (specific activity

of 10^6 cpm/µg DNA), using hybridization conditions determined earlier from Southern hybridization of the rat vasopressin gene (pBS 13+) to lobster genomic DNA (section 2.2.1.2, Chapter II). Plaques corresponding to the position of positive signals observed in the autoradiogram were each transferred into 1 ml of SM (0.1 M NaCl, 8.11 mM $MgSO_4.7H_20$, 50 mM Tris-HCl, pH 7.5, 0.01% gelatin) and stored at 4°C with a drop of chloroform.

Fifty six phage samples were collected and screened a second time at a density of 500 pfu per 90 mm plate. Out of this, 20 samples were isolated and screened a third time at a density of 250 pfu. The 947 PCR fragment was then used in the third round of screening as it was shown to be a more specific probe for use in the isolation of the putative MIH gene sequence. In Southern blot analysis of λ DNA prepared from plate lysate stocks of these clones (Sambrook *et al.*, 1989; Santos, 1991), the 947 PCR sequence hybridized to two *Sac*I restriction fragments from two genomic clones. The final post-hybridization wash used was increased to 0.1 x SSC/0.1% SDS at 65°C for 15 min, and the blot was exposed for autoradiography for 5 days.

(iii) Subcloning of lobster genomic inserts

*Sac*I digested fragments corresponding to the positions of the hybridization positive bands in these samples were gel purified using the Prep-A-Gene DNA purification kit (Bio-Rad), and cloned into the *Sac*I restriction site of the plasmid vector. Equimolar quantities of *Sac*I digested inserts were ligated with 100 ng of phosphorylated *Sac*I digested pBS m13+ vector (plasmid BlueScript m13+, Stratagene), according to Sambrook *et al.* (1989). The entire ligation mixture was used to transform DH5α competent cells (Maniatis *et al.*, 1982), and plated onto LB plates containing 50 µg/ml ampicillin, 20 µg/ml Xgal, 0.5 mM IPTG. White coloured bacterial colonies containing recombinant plasmids were transferred onto another agar plate and screened by colony hybridization to the 947 PCR fragment according to Sambrook *et al.* (1989).

Plasmid DNA was isolated from two samples, plasmid p2A and plasmid p66, using the WIZARD maxiPREP DNA kit (Promega). The two samples, plasmid p2A and plasmid p66, contain lobster genomic inserts of approximately 7.4 and 8.0 kb in length respectively. The presence of the 947 PCR sequence was subsequently confirmed by Southern hybridization of restriction digests of these plasmids, using a final wash stringency of 0.5 x SSC/0.1% SDS at 65° for 15 min.

(iv) Partial sequencing of the genomic clones

The genomic inserts in plasmids p2A and p66 were sequenced from both ends initially (200 nucleotide). To sequence further into the insert, the nested deletion strategy was used to obtain a set of nested deletion subclones containing progressive unidirectional deletions of the genomic insert, for both plasmids p2A and p66 (Sambrook *et al.*, 1989). Several sets of nested deletion clones were generated from plasmids p2A and p66 using the Erase-A-Base system (Promega).

Plasmids were sequenced from one end, T3 promoter primer end, using the ^{T7}Sequencing kit (Pharmacia LKB). Sequencing reactions were carried out on 2 µg of plasmid DNA, using 35 ng of T3 promoter primer, 10 µCi of [α-^{32}P] dCTP (3000 Ci/mmol); sequencing reactions were analyzed on a 6% polyacrylamide-8.32 M urea sequencing gel in TBE. Sequencing of plasmids was carried out to obtain overlapping sequences. To resolve some of the unclear regions of sequences, the sequence data from several clones was compared. The overlapping sequences were also confirmed in some cases by automated sequencing, which allowed approximately 500 nt to be read. DNA sequences were analyzed for overlapping sequences using the connect program, and homology searches on DNASIS HBIOI.

3.2.2 Isolation of the putative MIH from the lobster eyestalk cDNA library

3.2.2.1 Preparation of lobster eyestalk mRNA for cDNA library construction

The preparation of total RNA from lobster eyestalks, and the purification of mRNA for cDNA library construction has been described in section 2.2.3.3 (Chapter II). Prior to cDNA synthesis and cloning, the integrity of mRNA isolated from the

eyestalks was examined by determining the ability of the mRNAs to produce proteins in a wheat germ translation kit (Boehringer Mannheim). Each reaction contained approximately 1.5 µg of mRNA, 5 µl of translation reaction mixture, 125 mM potassium acetate, 125 mM magnesium acetate, 2 µl of L-[^{35}S] methionine (20 µCi, >1000 Ci/mmol), and 7.5 µl of extract, total volume of 25 µl. A water control, and 0.25 µg of globin mRNA were also run in parallel with each set of reactions. The mRNAs were heated at 65°C for 10 min and cooled on ice to disrupt any secondary structures in the mRNAs. The reactions were incubated at 30°C for 2 h, and the products stored at −80°C.

The labelled proteins produced in this assay were analysed by electrophoresis on 12% polyacrylamide gel containing 0.1% SDS, using a Bio-Rad PROTEAN IIx slab gel apparatus and running buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS, pH 8.3) (Laemmli, 1970). Five to ten µl of each sample was made up to 25 µl, in sample buffer (2% SDS, 10% glycerol, 0.06 M Tris-HCl, pH 6.8, 0.005% bromophenol blue, 0.1 M DTT). The samples were heated in a boiling water bath for 5 min and cooled on ice before loading. Samples were run at 150 V for 3 h 30 min, against high molecular weight protein markers (Bio-Rad).

After electrophoresis, the molecular weight markers were cut off and stained in 0.25% Coomassie blue R-250 dissolved in 50% methanol, 10% acetic acid for at least 30 min. The gel was destained by several washes in methanol:water:acetic acid (50:40:10; v:v:v), followed by 10% methanol:10% acetic acid solution. The portion of gel containing the radiolabelled samples was fixed for 30 min in isopropanol:water:acetic acid (25:65:10; v:v:v), and then incubated for 15-30 min in Amplify (Amersham). The gel was dried for 1 h at 80°C in a gel dryer, and exposed to X-ray film for 3 days.

3.2.2.2 Construction of cDNA library

The mRNA isolated was used as a template for cDNA synthesis, which was subsequently cloned into lambda gt10 (λ gt10), a bacteriophage vector capable of accepting inserts up to 7.6 kb in length (Huynh *et al.*, 1985). Lamda gt10 was chosen

as a vector for cDNA cloning, as the cDNA library constructed was to be screened using a DNA probe.

(i) Preparation of vector DNA, and vector arms for cloning

Detailed descriptions of the bacterial strains, cloning vectors and bacteriological methods are listed in Appendix 1.

Lambda gt10 DNA was prepared from plate lysate stocks according to the zinc chloride precipitation procedure of Santos (1991). Plate lysate stocks of λgt10 (10^{10}-10^{11} plaque forming units per ml, pfu/ml) were prepared according to protocol I in Sambrook *et al.* (1989), using *E. coli* strain LE 392 as the bacterial host for phage growth (Appendix 1). Amplified phages were harvested in λ diluent (10 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$).

Vector arms for cloning were prepared according to Jendrisak *et al.*, (1989) and Sambrook *et al.* (1987). The vector arms were prepared by *Eco*RI digestion (5 U/μg DNA) of vector DNA at a concentration of 160 μg/ml for 6-8 h at 37°C. The digested vector was purified with phenol:chloroform extraction, followed by ethanol precipitation, and dissolved in TE pH 8 at approximately 100 μg/ml. *Eco*RI digested vector arms were dephosphorylated with calf intestinal alkaline phophatase (CIAP, Boehringer Mannheim; 0.01 U/10 μg DNA) to reduce the background of non-recombinants derived from self-ligation of the vector (Sambrook *et al.*, 1989). A sample from each reaction step was removed, ligated, and packaged *in vitro* (Rosenberg, 1987), and the efficiency of each treatment was determined. Ligated dephosphorylated arms produced 10^5 pfu, which was 30-fold less pfu then intact vector DNA. This sample of arms was used in the trial ligations, whereas dephosphorylated arms which were 100-fold less efficient, were used in the scaled-up cloning experiments.

(ii) cDNA synthesis from eyestalk mRNA

cDNA was synthesized from 5 μg of poly(A)$^+$ RNA using the cDNA TimeSaver cDNA synthesis kit (Pharmacia). Five hundred ng of oligo$(dT)_{12\text{-}18}$ primer was used

in the reaction, and cDNA synthesis was monitored by the addition of 20 µCi of [α-^{32}P] dCTP (10 week old label). Following the second strand reaction, 10 µl was removed and analysed by electrophoresis on a 1% agarose gel, against nick translated radioactively labelled DNA molecular size markers. cDNA synthesized from lobster eyestalk mRNA ran as a smear between 0.46 kb to 2.3 kb on the agarose gel, with a distinct band at 1.3 kb.

*Eco*RI/*Not* I adaptors were ligated to the blunt ended cDNA according to the protocol, and the excess adaptors removed by spun column chromatography through Sephacryl S-400 gel.

(iii) cDNA cloning

Trial ligations of 15, 10 and 5 µl of column effluent, with 2 µg of *Eco* RI cut, dephosphorylated lambda gt10 vector, were set up as "precipitation mixtures". The DNA was then precipitated out of solution with 0.3 M sodium acetate and ethanol. Ligation was carried out in 10 µl at 16°C for 16 h. A fraction of each reaction was packaged *in vitro*, using packaging extract prepared from *E. coli* strain SMR 10 (Rosenberg, 1987), and the packaged plaques were titrated on TB plates. Trial ligations of insert to vector at 15, 10 and 5 ng of cDNA to 2 µg of vector DNA gave library sizes between 10^4 to 10^6 pfu/µg of cDNA.

Serial dilutions of each packaged samples were made, and 0.1 ml of each dilution was added to 0.1 ml of LE 392 plating cells (Appendix 1). The plaques were allowed to adsorb at 37°C for 20 min, and the cells plated in molten top agarose overlayed on 0-2 day old TB plates. Plates were incubated upside down at 37°C overnight. The titre of the packaged phage was determined by counting the number of plaques formed on the bacterial lawn. Recombinant λgt10 phages formed clear plaques, whereas, non-recombinant phages formed turbid plaques on LE 392 host cells. The titre of packaged plaques was calculated by the following equations:

Titre (pfu/ml) =(No. of plaques x dilution factor)/amount plated in ml.

Titre (pfu/µg DNA) = Titre/(DNA packaged in 0.5 ml x 2).

The amount of DNA in each reaction was determined by comparison of 1 µl of DNA samples with lambda DNA standards spotted onto the surface of a 1% agarose gel containing 0.5 µg/ml ethidium bromide, and visualized over a UV transilluminator (Sambrook *et al.*, 1989).

Using the proportions that gave the largest number of recombinants (pfu/µg cDNA) (15 ng insert : 2 µg vector) scaled up ligations were performed, and the entire ligation reaction was packaged, and titred. Several cDNA libraries were constructed, ranging in size from 10^4-10^5 pfu/µg of cDNA, with a combined total of 5.6 x 10^6 pfu/µg of cDNA. The libraries generated in the trial, and scaled up reactions were then screened by plaque *in situ* hybridization according to Davis and Benton (1977) using the 947 PCR fragment initially and, subsequently, with a genomic clone isolated from the lobster genomic DNA library (section 3.2.1) in conjunction with the 947 PCR fragment.

3.2.2.3 Screening of cDNA libraries for the putative MIH sequence

Initial screening of the entire cDNA library with the 947 PCR sequence failed to retrieve any definitive clones after three rounds of plaque purification. Standard library screening procedures as described in Sambrook *et al.* (1989) were followed. The post-hybridization washes were carried out at 65°C, in 2 x SSC for 2 x 15 min and finally 2 x SSC/0.1% SDS for 30 min. Filters were exposed to X-ray film for 3-5 days.

After the primary screen, plated out cDNA libraries were recovered by overlaying the plates with 5 or 12 ml of SM (90 or 150 mm diameter plate), and allowing the phages to elute out overnight, at 4°C. The bacteriophage suspension was recovered, and the plate rinsed with 1-4 ml of SM. The suspension was incubated with 0.2 ml of chloroform for 15 min, at room temperature with occasional shaking. Bacterial cell and agar debri were removed by centrifugation at 4,000 x *g* for 5 min at 4°C. The amplified libraries were stored in dark glass bottles at 4°C with a drop of chloroform.

Failure to isolate the putative MIH gene sequence from the cDNA library may have been due to the fact that the 947 PCR fragment was derived from genomic DNA, and hence the protein coding sequence (exon) could be represented in only a small proportion of the fragment. Screening of the amplified cDNA library was repeated using a combination of several methods, as well as a longer probe containing the 947 PCR sequence.

Based on screening of the lobster genomic DNA libraries with the 947 PCR product (section 3.2.1), two genomic DNA fragments, 7.4 and 8.0 kb, were isolated and these were subcloned into plasmid Bluescript pBS m13+ (Stratagene) as clones p2A, and p66. Plasmid p66 was used as a probe in the initial screen of the amplified cDNA libraries. Clones isolated from the primary screen were subsequently screened with the 947 PCR fragment.

Plasmid p66 was used as a probe in the initial screening of the amplified cDNA libraries. Approximately 300,000 pfu were screened; these were plated at a density of 50,000 pfu per 150 mm plate (TB agar). Plaques were blotted onto nylon filter circles (Hybond-N, Amersham), as described in Sambrook *et al.* (1989), and by Amersham. The filters were then incubated for 7 min on a pad of 3MM paper soaked in denaturation solution (1.5 M NaCl, 0.5 M NaOH), followed by 2 x 3 min in neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl, pH 7.2, 1 mM EDTA), and rinsed finally in 2 x SSC. Once dry, DNA was cross-linked to the nylon filters by placing the filters face down on a standard UV transilluminator for 3 min.

Up to 6 or 20 filters (150 mm or 90 mm diameter) were hybridized in 20-25 ml of prehybridization solution (6 x SSC, 5 x Denhardt's, 0.5% SDS, 0.5% blocking agent (Amersham)) containing 25 ng of denatured, [α-^{32}P] dCTP random prime labelled 947 PCR fragment (specific activity of 10^8-10^9 cpm/μg), at 65°C for 16 h. Post-hybridization washes were carried out at 65°C in 2 x SSC for 2 x 15 min, 2 x SSC/0.1% SDS for 30 min, and finally in 1 x SSC/0.1% SDS for 15 min. Filters were exposed to X-ray film for 3-5 days.

Between 4-10 plugs of agar were removed into 1 ml of SM using a Pasteur pipette tip, from areas on the plate showing positive hybridization to plasmid p66, on the corresponding filter. The phages were eluted out at 4°C, for several hours. Phage DNA was prepared from plate lysate stocks of the eluted phage samples (Santos, 1991; Sambrook *et al.*, 1989). The phage DNA was first analysed by PCR, and then by Southern blotting of phage DNA and PCR products using the 947 PCR fragment as a probe.

Klein *et al.* (1993b) described a technique for screening phage libraries using PCR, in combination with plaque *in situ* hybridization to isolate the clone of interest. Based on this, PCR was performed on phage DNA samples according to methods described in section 2.2.2.1 (Chapter II), at the annealing temperature of 55°C, using 2.5 mM $MgCl_2$, 1.4 U *Taq* polymerase, and internal primers designed from the 947 PCR sequence:

V_{int1} 5'–GTAAGTACATCGCTTG–3', and

V_{int2} 5'–TCCAGAAGTCGGTTTAAG–3'.

Reaction tubes were held at 95°C for 10 min, before adding the *Taq* enzyme at 80°C. Half of the PCR reaction was then analysed on a 1% agarose gel, but no specific bands were observed. The remainder of the PCR products were then analysed, together with approximately 1 µg of each phage DNA sample, by Southern hybridization to the 947 PCR fragment as described in section 2.2.3.1 (Chapter II). A final post-hybridization wash of 0.5 x SSC/0.1% SDS at 65°C for 10 min was used, and the membranes were exposed for autoradiography for 1-4 days.

Samples of eluted phages which gave a positive signal when hybridized with the 947 PCR probe were then plated at a lower density of 500 pfu per 90 mm dish, and screened by *in situ* plaque hybridization, with a final wash at 1 x SSC/0.1% SDS at 65°C for 10 min. The plaques hybridizing to the probe were well separated on the plate, and were picked as single plaques into SM. The phage isolates were titred, and then plated at 100-200 pfu per plate, for the final screening with the 947 PCR fragment.

3.2.2.4 Subcloning of cDNA clones:

(i) Isolation of cDNA inserts

Purified cDNA clones showing positive hybridization to the 947 PCR fragment were then selected for sequence analysis. The cDNA inserts were excised out and subcloned into a plasmid expression vector for sequence analysis, as well as expression of the protein. Phage DNA was extracted from plate lysate stocks of cDNA clones using the WIZARD prep kit from Promega. cDNA inserts were excised from recombinant λgt10 DNA using *Not*I restriction endonuclease (New England Biolab), as there was an internal *Eco*RI site within some of the cDNA inserts. *Eco* RI is the cloning site of the vector, whereas *Not* I is an internal restriction site found in the adaptors ligated to the blunt-ended cDNA prior to cloning. The excised cDNA inserts were then gel purified on a 0.7% agarose gel in 1 x TAE, and the DNA extracted using the Prep-A-Gene DNA purification kit (Bio-Rad).

(ii) Preparation of the vector

The cDNA inserts were cloned into the *Not*I restriction site of a cDNA cloning plasmid vector, pSPORT 1 (obtained from Life Technologies, GIBCO). Fifty nanograms of plasmid DNA was used in transforming *E. coli* strain DH5α competent cells (Maniatis *et al.*, 1982). Plasmid DNA was then prepared using the plasmid maxi-kit according to the manufacturer's instructions (QIAGEN).

Plasmid DNA was digested in a 3-fold excess of *Not*I at 37°C for 4-8 h. The digested DNA was purified by phenol:chloroform, and chloroform extraction, and ethanol precipitated with 0.3 M sodium acetate, pH 7.0, at –20°C, overnight. The recovered digested plasmid was then dephosphorylated with a 5-fold excess of calf intestinal alkaline phophatase (1 U/10 pmoles ends) (CIAP, Boehringer Mannheim), at 37°C for 60 min. The desphosphorylated plasmid was purified by phenol:chloroform and chloroform extraction, and ethanol precipitated.

(iii) Ligation of insert to vector

Equimolar quantities of plasmid DNA and cDNA insert were ligated with 0.4 Weiss units of T4 DNA ligase and 0.5 mM ATP, in ligation buffer (50 mM Tris-

HCl, pH 7.6, 10 mM $MgCl_2$, 20 mM DTT), at 16°C for 12-16 h, according to Sambrook *et al.* (1989). The entire ligation mixture was then used to transform DH5α competent cells and plated on LB agar plates containing 50 µg/ml ampicillin, 0.5 mM IPTG, 20 µg/ml X-gal.

Plasmid DNA was prepared from 1.5 ml overnight cultures of several white bacterial colonies of each transformation reaction according to the boiling preparation method described in Sambrook *et al.*, 1989, which was adapted from Holmes and Quigley (1981). The DNA was stored in 50 µl of TE containing DNase-free pancreatic RNase (20 µg/ml), at -20°C. The clones were then tested for the presence of cDNA insert by restriction with *Not*I enzyme and analyzed by agarose gel electrophoresis.

(iv) Sequencing of cDNA inserts

Double stranded plasmid DNA templates, isolated by the boiling method, were sequenced by the dideoxy sequencing method (Sanger *et al.*, 1977). Plasmids were sequenced from both ends using the universal primer, and the T7 promoter primer (Promega). Sequencing reactions were carried out according to the manufacturer's instructions, on 2 µg of plasmid DNA template using 30-35 nanograms of primer and 10 µCi of [α-^{32}P] dCTP (3000 Ci/mmol) per reaction, using the ^{T7} Sequencing kit (Pharmacia LKB).

Two microliters of each sample were analysed on a model S2 sequencing gel apparatus (BRL), using 6% polyacylamide-8.32 M urea sequencing gel in TBE. As the primer sites were between 61-74 nucleotides from the insert, sequencing reactions were run (constant voltage of 1600 V, 60 W (max)) for 6 h in the long run, and 3 h for the short run (1 hour after the bromophenol blue left the edge of the gel).

The gel was then transferred onto a Whatman 3MM paper, and dried under vacuum at 80°C for 30 min (sequencing cycle) on a gel dryer (BIO-RAD). The dried gel was exposed without intensifying screens, to Kodak X-Omat AR film for 6-18 h at −80°C. With the conditions used, it was possible to read between 280-340 nt from

each end of the insert. The remainder of the sequence data was obtained from automated sequencing of the template from one end, to obtain the sequence that overlaps between the two ends. Plasmid DNA was extracted using a modified alkaline-lysis/PEG procedure (Applied Biosystems, Inc.), and the templates were sequenced at the Centre for Gene Research, University of Otago, Dunedin, New Zealand.

3.2.2.5 Sequence analysis of cDNA clones

The nucleotide sequences obtained were analysed on the computer using DNASIS, and searched for homology to other known sequences using the NCBI BLAST tool.

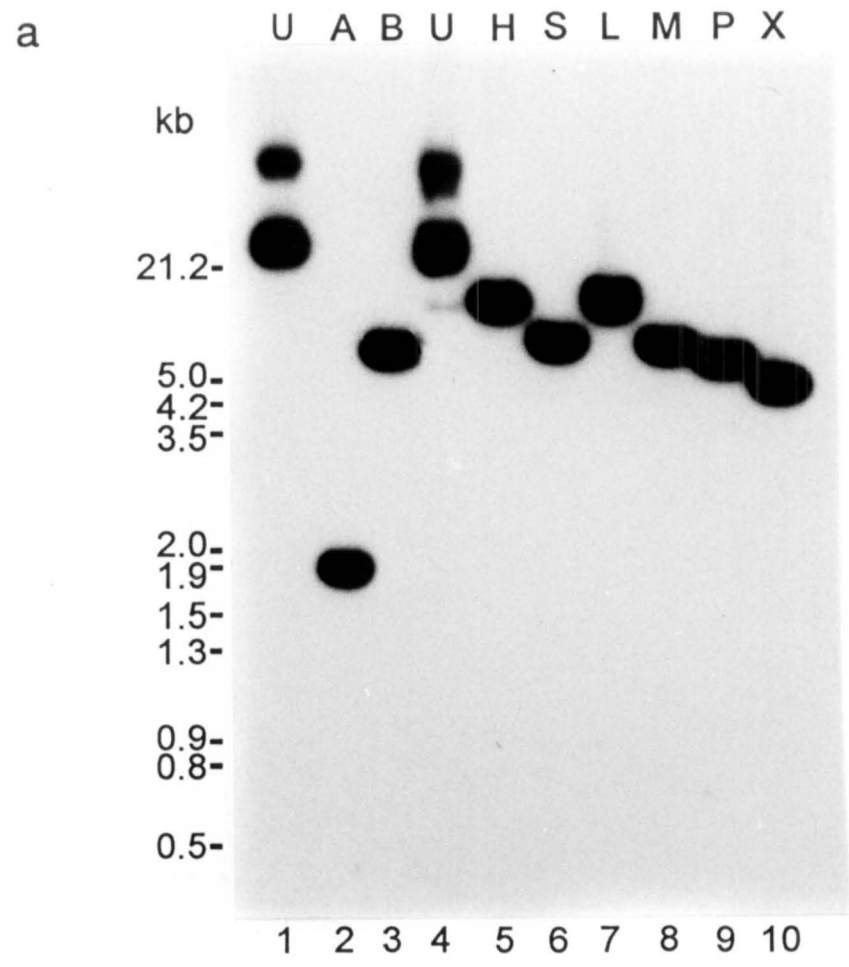
3.3 RESULTS

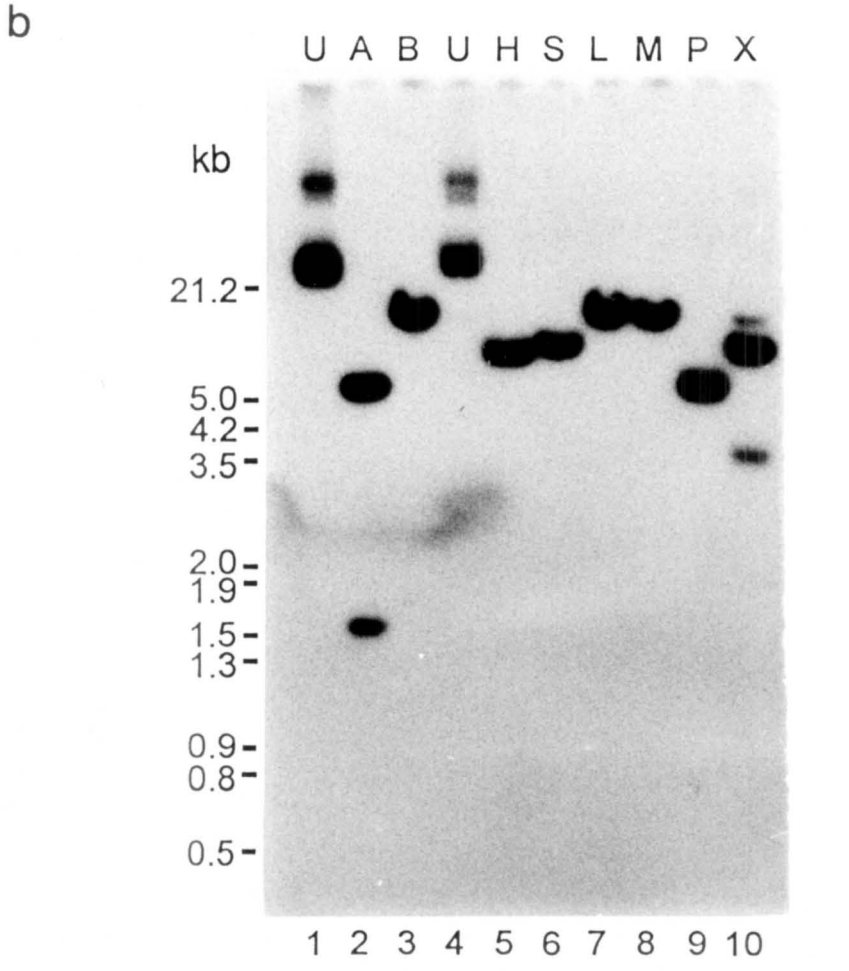
3.3.1 Isolation of the genomic clones containing the putative MIH gene sequence

As a result of screening the lobster genomic libraries with the 947 PCR product, two genomic clones were isolated. The SacI digested genomic inserts were subcloned into the plasmid vector, pBS m13+. Recombinant plasmids detected by the 947 PCR sequence were selected by colony hybridization, and the clones giving the strongest hybridization signal were isolated and sequenced. Two clones, p2A and p66, containing inserts of 7.4 and 8.0 kb were selected for further analysis.

Plasmids p2A and p66 were digested with several restriction enzymes, to look for cloning sites which were absent from the genomic inserts. In Southern blot analysis of these digests, the 947 PCR product hybridized to two *Acc*I restriction fragments (5.7 and 1.6 kb,) and three *Xba*I fragments (12.3, 7.4, 3.5 kb) within plasmid p66 (Fig. 3.1b, lanes 2 and 10). Based on its size and signal, the 12.3 kb band in the XbaI digest of plasmid p66 is linearized DNA. Hybridization of the 947 PCR product to single bands was observed for plasmid 2A (Fig. 3.1a).

Figure 3.1. Southern blot analysis of genomic clones, plasmids p2a and p66. Between 200-250 ng of plasmid DNA was digested with *Acc*I (A), *Hin*dIII (H), *Sac*I (S), *Sal*I (L), *Sma*I (M), *Sph*I (P), or *Xba*I (X). Undigested plasmids (U) were also included in the experiment. Restriction digests of plasmids p2A (**a**) and p66 (**b**) were hybridized to the 947 PCR product, and a final post-hybridization wash of 0.5 x SSC/0.1% SDS at 65°C for 15 min was used. The blots were exposed for autoradiography for 22 h at −70°C. The location of DNA size markers (*Eco*RI and *Hin*dIII digested λ DNA) are indicated (kb): 21.2, 21,226 bp; 5.0, 5,060 bp; 4.2, 4,268 bp; 3.5, 3,530 bp; 2.0, 2,027 bp ; 1.9, 1,904 bp; 1.5, 1,584 bp; 1.3, 1,375 bp; 0.9, 947 bp; 0.8, 831 bp; 0.5, 564 bp.


a
U A B U H S L M P X
kb
21.2
5.0
4.2
3.5
2.0
1.9
1.5
1.3
0.9
0.8
0.5
1 2 3 4 5 6 7 8 9 10


b
U A B U H S L M P X
kb
21.2
5.0
4.2
3.5
2.0
1.9
1.5
1.3
0.9
0.8
0.5
1 2 3 4 5 6 7 8 9 10

A comparison of the number and relative mobilities of these bands in the ethidium bromide stained gel (figures not shown) and Southern hybridization analysis (Table 3.1) suggested that the insert in plasmid 2A was different from that in plasmid p66. Molecular sizes of the restriction bands were estimated by comparison to *Eco*RI and *Hin*dIII digested λ DNA markers run on the gel. The band sizes listed in Table 3.1 are only estimates; the size discrimination was limited by the range of separation on the 10 cm minigel.

Table 3.1. Estimates of band sizes obtained for plasmid p2A and p66, when cut with the restriction enzymes which also cut in the plasmid polylinker sequence. Bands showing hybridization to the 947 PCR product (Fig. 5.2a and 5.2b) are underlined. The molecular sizes were determined by comparison to both *Hin*d III and *Eco*RI + *Hin*d III digested λ DNA markers.

Enymes	Plasmid p2A		Plasmid p66	
	No. of bands	Band size (kb)	No. of bands	Band size (kb)
*Acc*I (A)	5	3.5, 2.5, 1.9, 1.6, 0.7	5	5.7, 2.6, 1.6, 1.1, 0.7
*Bam*HI (B)	2	7.0, 3.9	1	12.3
*Hin*dIII (H)	1	10.8	2	7.4, 4.2
*Sac*I (S)	2	7.4, 3.2	2	8.3, 3.3
*Sal*I (L)	1	10.8	1	12.3
*Sma*I (M)	2	7.4, 3.2	1	12.3
*Sph*I (P)	2	6.5, 4.2	2	6.1, 5.4
*Xba*I (X)	3	5.0, 3.2, 1.7,	3	12.3, 7.4, 3.5

A search for these restriction sites (restriction sites found within the plasmid polylinker sequence) within the 947 PCR sequence showed that none of the restriction sites listed in Table 3.1 were present within the 947 PCR sequence. However, the 947 PCR product hybridized to more than one *Acc*I and *Xba*I restriction fragment in plasmid p66. This suggested that either the sequences detected in plasmid p66 were not the actual 947 PCR sequence, but closely related to it, or, there was more than one 947 PCR related sequence in the genomic insert in plasmid

p66. Medium stringency conditions (0.5 x SSC/0.1% SDS at 65°C for 15 min) and 200 ng (plasmid p2A) and 250 ng (plasmid p66) of plasmid DNA were used in the Southern blot analysis. The relatively short exposure time of 22 h and the strong hybridization signals obtained suggested that the sequences detected by the 947 PCR product (specific activity 10^6 cpm/µg) were highly related if not the 947 PCR sequence itself. Based on this result plasmid p66 was used as probe to screen the cDNA library for the putative MIH gene sequence.

3.3.2 Sequence analysis of plasmids p2A and p66

The sequenced regions within the genomic inserts in plasmids p2A and p66, are indicated in Fig. 3.2 as 2A-1 to 2A-5, and 66-1 to 66-3, respectively. The relative position of the 947 PCR sequence in plasmids p2A and p66 was determined by DNA dot blot analysis (medium stringency) of the nested deletion clones generated for sequencing. Hybridization to the 947 PCR sequence was absent in plasmids less than 5.3 kb and 5.0 kb in size for p2A and p66 respectively, after 22 h exposure (data not shown), suggesting that the 947 PCR related sequence was found in the deleted regions of these plasmids, as indicated in Fig. 3.2.

However, a search using DNASIS failed to locate the 947 PCR sequence within any of the sequenced regions of plasmids p2A and p66 section A2.2 (Appendix 2). Each sequence was individually searched for homology to the 947 PCR sequence, and also to peJK1 and peJK2 cDNA sequences, in both normal and complementary strand orientation, using DNASIS. As sequencing of the plasmids was not completed, the sequence data are presented in Appendix 2 section A2.3 (Appendix 2).

3.3.3 Isolation of the putative MIH cDNA clones

3.3.3.1 Analysis of mRNA

The appearance of total RNA and poly(A)$^+$ RNA isolated from lobster eyestalks was shown in Fig. 2.10a (Chapter II). The integrity of mRNA purified from lobster eyestalks was determined by its ability to direct protein synthesis *in vitro* in wheat germ extract (Fig. 3.3). Translation of poly(A)$^+$ RNA from lobster eyestalk produced >20 different radioactively labelled protein bands ranging in size

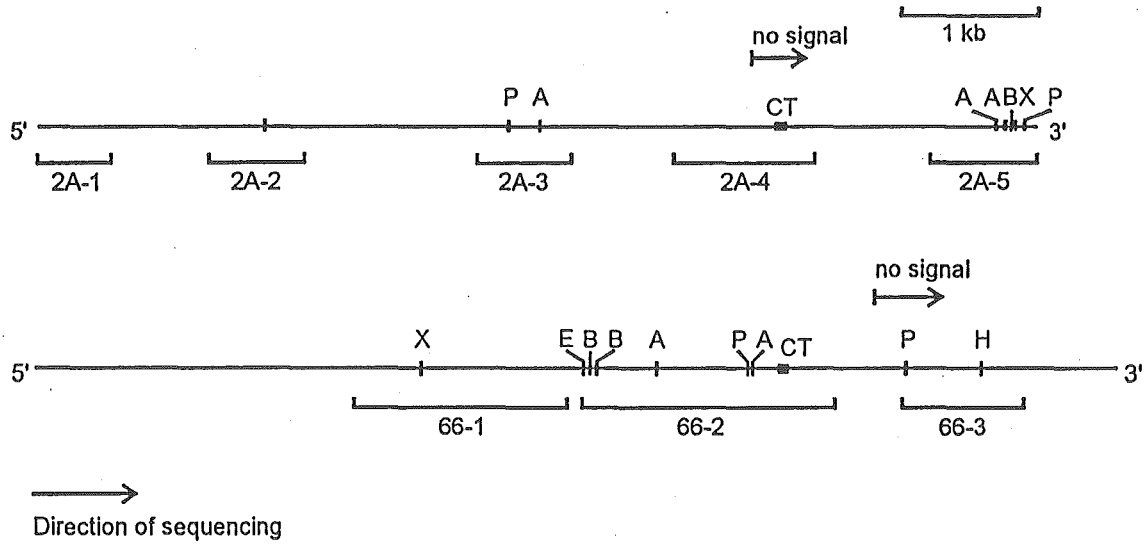
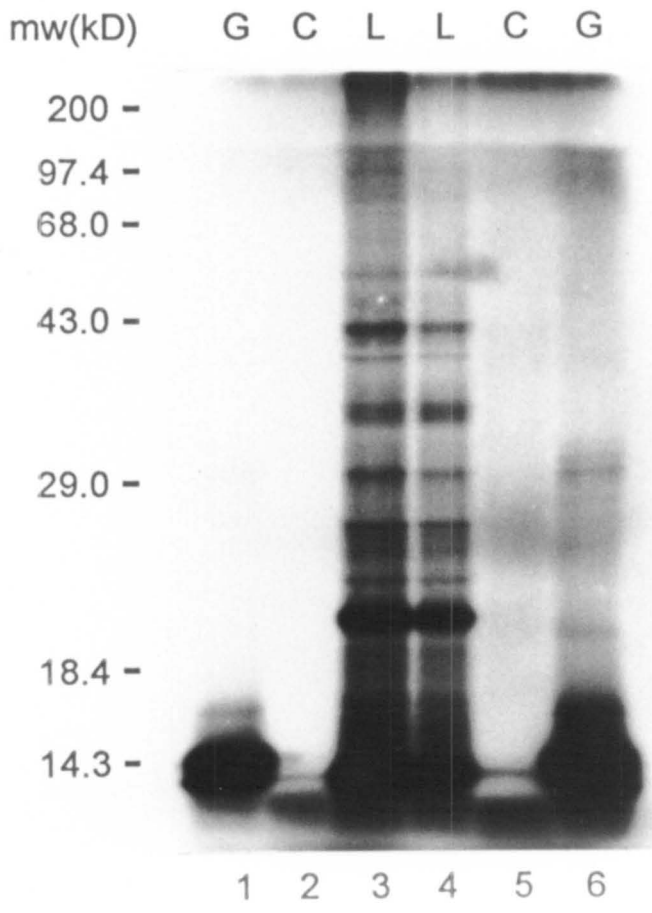


Figure 3.2. Schematic diagram of lobster genomic clones p2A (7.4 kb) and p66 (8.0 kb). The sequenced regions (2A-1 to 2a-5, and 66-1 to 66-3) are indicated by brackets. Restriction sites of *Acc*I (A), *Bgl*II (B), *Eco*RI (E), *Hin*dIII (H), *Sph*I (P), and *Xba*I (X) are indicated; CT represents the CT repeat sequence in 2A-4 (70 bp) and in 66-2 (54 bp). Hybridization to the 947 PCR sequence was abolished when the 5′ region of the clones were deleted past the arrow (no signal), thus indicating that the location of the 947 PCR sequence is upstream from the arrow.

Figure 3.3. Fluorogram of L-[^{35}S] methionine labelled proteins from wheat germ cell free translation assay of *Xenopus* β-globin mRNA (G), lobster eyestalk poly(A)$^+$ RNA (L), and water control (C). Five microlitres of translation product was loaded in lanes 1, 2, and 4; and 10 µl in lanes 3, 5 and 6, on a 12% polyacrylamide SDS gel. Protein molecular weight standards are indicated on the left. The dried gel was exposed for fluorography for 3 days at −80°C, with intensifying screens.


mw(kD)
G C L L C G
200
97.4
68.0
43.0
29.0
18.4
14.3
1 2 3 4 5 6

from 13.8 kD to 93.6 kD, with two major bands at 13.8 kD and 20.7 kD (Fig. 3.3, lanes 3 and 4). β globin mRNA, included as a positive control in the reaction produced at least four protein bands ranging in size from 29.9-13.6 kD, with a major band at the expected molecular weight of 16.6 kD (lanes 1 and 6). Discrimination of the major band was not clear due to overexposure, and the separation of the gel for that size range. There was no protein band produced in the water control.

3.3.3.2 Screening of amplified cDNA libraries with the plasmid p66 (genomic library clone), followed by the 947 PCR product

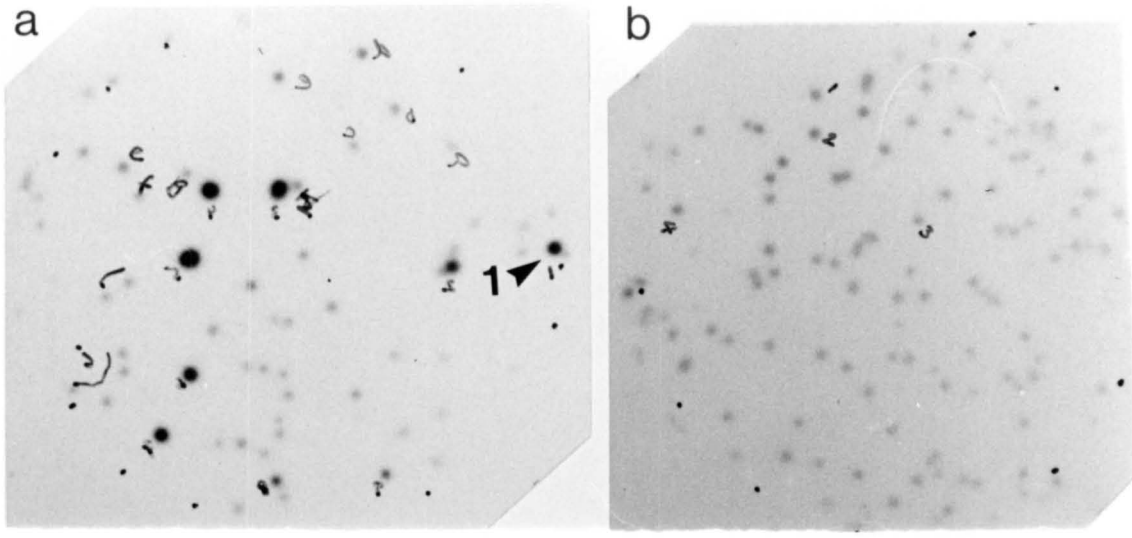
This approach was taken as it was postulated that as the 947 PCR fragment isolated from lobster genomic DNA contained both exon and intron sequences, the chances of detecting a cDNA clone could be increased by using a larger probe such as plasmid 66, isolated from the genomic library of lobster using the 947 PCR product as a probe.

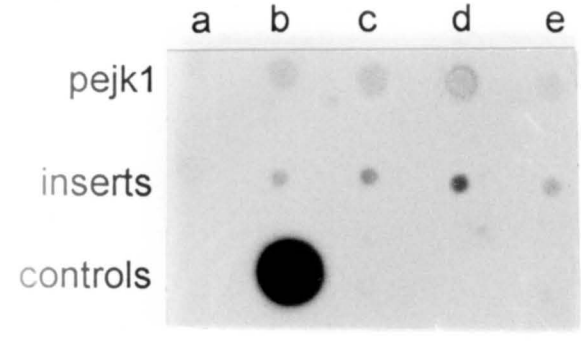
In the initial screening of the recovered cDNA libraries with plasmid 66, 29 samples of plaques were removed from areas on the plates showing positive hybridization to the probe. The hybridization signal was weak. Analysis of the DNA isolated from these plaque samples, by PCR, failed to detect any 947 PCR product. The PCR primers were designed from internal sequences within the 947 PCR product. Based on sequence analysis of the 947 PCR product, these sequences were located outside the predicted protein coding region. Ten of these samples (1 µg DNA), however, showed positive signal when probed with the 947 bp PCR product in a Southern hybridization analysis.

In the secondary screening of positive samples with the 947 bp PCR product, the signals obtained were very definitive (Fig. 3.4a), corresponding to individual plaques on the plates, and were just as strong for both the nylon and nitrocellulose filters. The clones showing the most promising hybridization signals were then plated (100-200 pfu) and hybridized to the 947 bp PCR fragment. Of the samples tested, 15 plates showed positive hybridization in all of the plaques except for some at the edges, indicating that the plaques were purified sufficiently (Fig. 3.4b). The intensity

Figure 3.4. cDNA library screening by plaque *in situ* hybridization to the 947 PCR probe. (**a**) Secondary screening of sample 3C1, plated at 500 pfu per plate and (**b**) final sceening of sample 3C1-1, clone **1** (▸) in (a), plated at 150 pfu per plate. Phages were grown at 37°C for 8 h. Nylon filters were washed in 1 x SSC/0.1% SDS at 65°C for 10 min, and exposed for autoradiography for 4 days.

Figure 3.5. DNA dot blot hybridization to the 947 PCR fragment. Plasmid recombinant clones peJK1, 200 ng (a-e), and cDNA inserts (200 ng) excised from λgt10 recombinant phages (b, d, e, cloned as peJK2, and c, as peJK3). Controls included were 8 ng 947 PCR product (b), and 200 ng of DNA from vectors; (a) λgt10, (c) *Sma*I digested pBS m13+, and (e) *Not*I digested pSPORT. The filter was washed in 0.5 x SSC/0.1% SDS at 65°C for 10 min, and exposed for autoradiography for 45 h.


a
b


a
b
c
d
e
pejk1
inserts
controls

of the signal however, was weaker than that in the previous screening, and also varied with the sample. Four separate plaques were picked from each plate, and DNA was prepared from plate lysate stocks derived from individual plaques.

3.3.4 Subcloning of cDNA library clones

The cDNA inserts that were excised out of the λgt10 vector were approximately 600 bp in length for all 15 clones. DNA dot blot analysis of the cDNA inserts, subcloned into the *Not*I restriction site of pSPORT 1, showed positive hybridization to the 947 PCR fragment after medium stringency wash (0.5 x SSC/0.1% SDS at 65°C for 10 min). The hybridization signal, after 18 h exposure, was weak for the pSPORT clones, and darker for the actual cDNA insert compared to the 947 PCR positive control (Fig. 3.5). The hybridization signal obtained was weaker for the plasmids compared to the cDNA inserts, as the plasmid clones were applied on the membrane in the intact state. Supercoiled plasmid DNA has to be converted to open circular, or linear form to bind to the membrane as denatured supercoiled DNA renatures too quickly on neutralization to be trapped in the denatured state on the membrane (Anderson and Young, 1985).

3.3.5 Sequence analysis of the cDNA clones

Each cDNA clone was partially sequenced from at least one end (200 nucleotides in) and, of the 15 clones isolated, there were actually only three different sequences. These were designated the names of peJK1 (10 out of 15 clones), peJK2 (4/15) and peJK3 (1/15).

The nucleotide and deduced amino acid sequences of the 585 bp cDNA cloned in peJK1 are shown in Fig. 3.6. Plasmid peJK1 contained an open reading frame with the initiation codon at nucleotide (nt) 133, encoding a 44 amino acid long polypeptide with a predicted molecular weight of 4813 daltons. The translation stop codon at nt 265 was followed by an untranslated region that ends with a poly(A) tail. A possible polyadenylation signal, AATATA (Birnstiel *et al.*, 1985), between nucleotides 548 and 555 in the untranslated region, was located 20 nucleotides upstream from the poly(A) tail.

```
  1 CACGAAGCGTGTGTAGTGCACAATCAGTTCCTACCACCACTACCACCTCTACCACCACCTACCACC

 67 TACCACCTACCACCACCACCACCACCACCTTCATCGTCATCTTCGCTGTCGTCGCCTCAAGACCTG

135 ATG CCC CCG GAG CCA CAG CTC GTC AGG AAG GTC CTC CCA GCA CAG GGA AGG
    Met Pro Pro Glu Pro Gln Leu Val Arg Lys Val Leu Pro Ala Gln Gly Arg

184 TCT TCA AAG GCC TCA TCC GTC ATC ATC ACC ACA GTC ATT CAC ATG TTA ACA
    Ser Ser Lys Ala Ser Ser Val Ile Ile Thr Thr Val Ile His Met Leu Thr

235 AGT CAA CAA CAA ACC CAG ATG GCC CAT TCA TGA CACTTTCTTTGGTTCTGCAGCTT
    Ser Gln Gln Gln Thr Gln Met Ala His Ser stop

291 TGAAGATTTGTTTCTGTGATATTATTTCTCTCTTTTGTATATCCATTTTTGGTTTTCTCATACGTTG

357 ACTGGGTTGGTCTGATTCAGAAGTGTAGTGAAGATAAGGATATATATGAACTTGACTTTAGGATTCT

425 ATTGAGTTGGTTCTGTTATTCATAAATGTTCTTCTCACTCTAAGTTCCTCTCGTCTGTTCTTGCCAT

492 CATTATCTCATGCTCACGAGGTATTTTTCTGTGGACATTAGAATGTATATTGGCCAAAATATATGAA

559 AGAATTGTATAAAAAAAAAAAAAAAA 3′
```

Figure 3.6. Nucleotide and deduced amino acid sequence of cDNA clone peJK1, 585 bp long, isolated from the lobster eyestalk. Both the initiation and stop codon are underlined, and the possible polyadenylation signal is in **bold-face** type.

cDNA clones peJK2 and peJK3 shared 96.6% homology over a stretch of 558 bp overlap. The nucleotide and deduced amino acid sequences of the 568 bp cDNAs from the two clones are shown in Fig. 3.7. Plasmid peJK2 contained an open reading frame encoding a 110 amino acid polypeptide with a predicted molecular weight of 11,177 daltons, a translation stop codon at nt 401 followed by an untranslated region which ended with a poly(A) tail 146 nucleotides downstream. The polyadenylation signal AATAAA was located between nucleotides 523 and 530, 26 nucleotides upstream from the poly(A) tail (19 residues long).

The nucleotide sequence of peJK3 was very similar to peJK2, containing an open reading frame encoding a 113 amino acid polypeptide of 11,602 daltons predicted molecular weight. The 5′ end of peJK3 begins 10 nucleotides upstream from peJK2, and except for 10 base substitutions (at nucleotides 131, 249, 273, 282, 313, 359, 504, 519, 550 and 564), and 9 additional bases (at nucleotides 143 to 145, and 169 to 174), the two sequences were essentially the same. The base substitutions cause a change in the amino acids encoded, and the additional bases code for an extra alanine and two valine amino acid residues. The translational stop codon at nucleotide 420 was followed by an untranslated region, with the AATAAA polyadenylation signal between nucleotides 542 and 549; however, the poly(A) tail is absent in peJK3. The poly(A) tail was probably lost during the cloning process as this feature was used to isolate the poly(A)$^+$ RNA initially, and also for binding of the oligo(dT)$_{12\text{-}18}$ primer in cDNA synthesis.

In the amino acid sequences encoded by peJK2 and peJK3, there was a large number of valine residues (25%). The hydrophobic profiles of the deduced amino acid encoded by the cDNA clones was performed using DNAMAN program, and the hydrophobicity plots of peJK1, peJK2 and peJK3 are shown in Figures 3.8a-c.

3.3.6 Sequence homology search

Nucleotide and deduced amino acid sequences of peJK1, peJK2 and peJK3 were compared to published nucleotide and protein sequences using DNASIS, BLAST, and FASTA programs. In the following sections the terms sequence

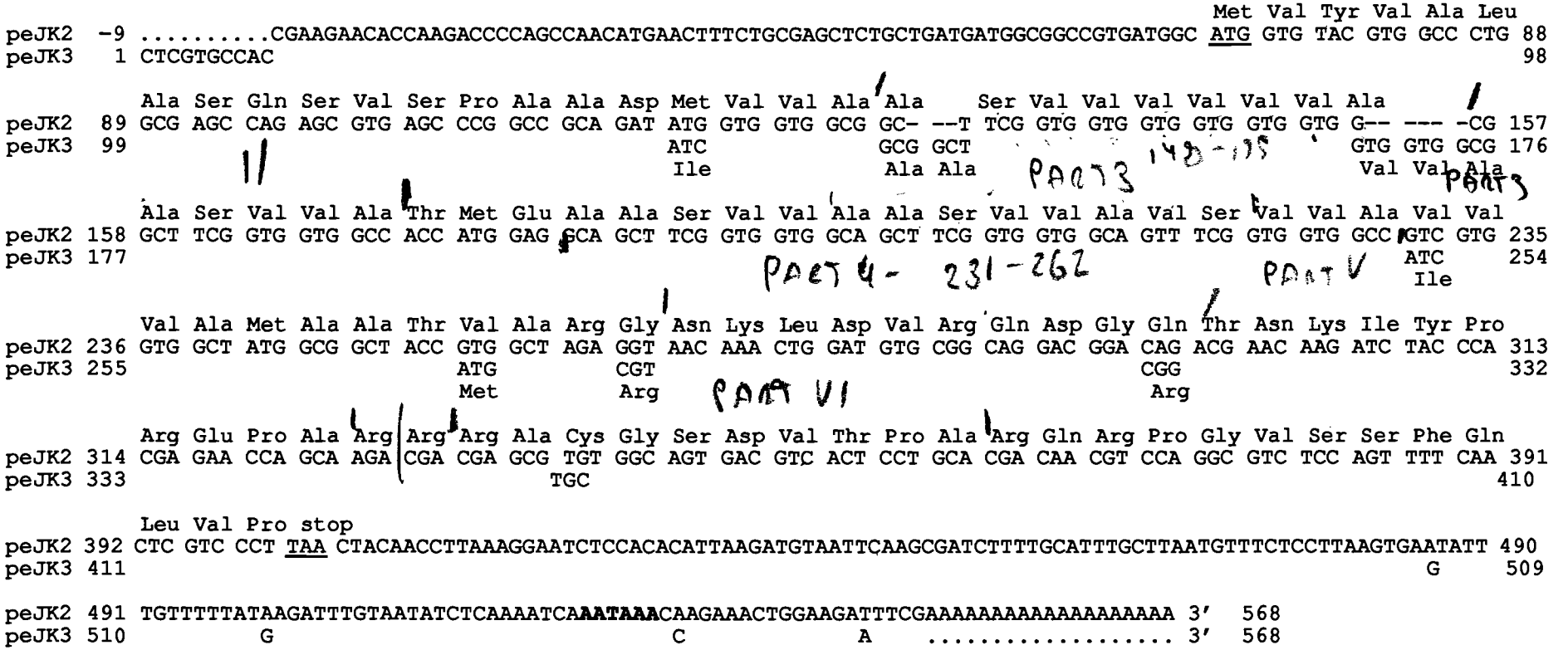


Figure 3.7. Sequence comparison of the nucleotide sequences and deduced amino acid sequences of cDNA clones peJK2 and peJK3 (568 bp long). The two clones were isolated from a cDNA library of the lobster eyestalk, and share 96% homology. The deduced amino acid sequence is shown above the nucleotide sequence; both the initiation and stop codons are underlined, and the polyadenylation signal **AATAAA** is indicated in **bold-face** type.

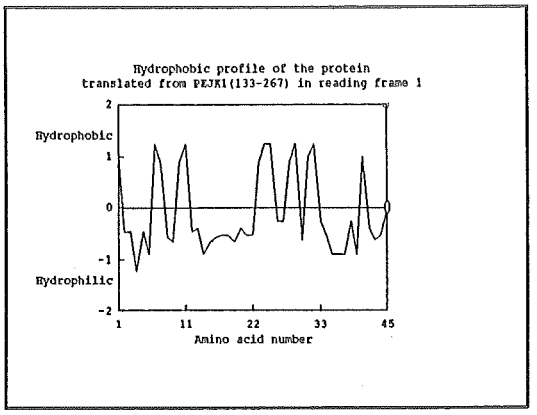


Figure 3.8a. Hydrophobicity plot of deduced amino acid sequence encoded by cDNA clone peJK1.

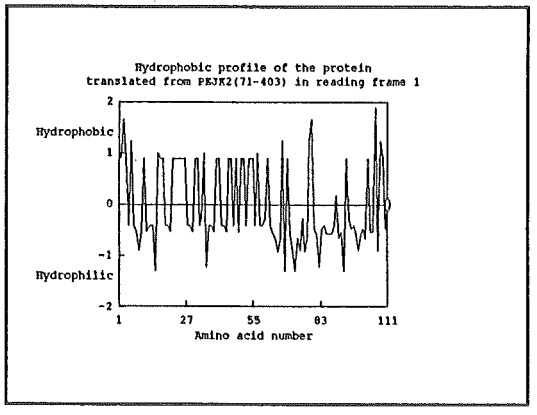


Figure 3.8b. Hydrophobicity plot of deduced amino acid sequence encoded by cDNA clone peJK2.

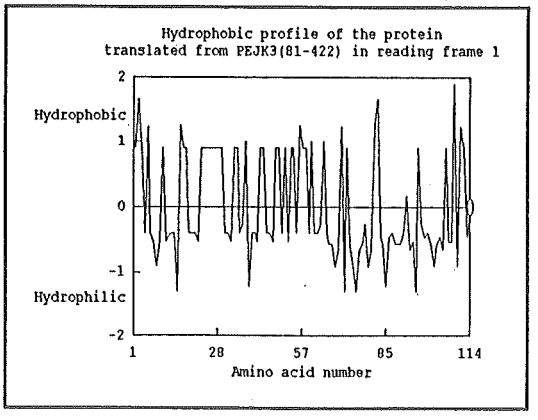


Figure 3.8c. Hydrophobicity plot of deduced amino acid sequence encoded by cDNA clone peJK3

homology, sequence identity and sequence similarity have been used interchangeably. However, the term homology implies that two sequences are related evolutionarily (von Heijne, 1987). In the nucleotide sequence homology searches using the DNASIS program, peJK1, peJK2 and peJK3 appeared to show some homology to known MIH and CHH cDNA sequences (Table 3.2).

The cDNA clones isolated in this study showed higher homology to the MIH sequence isolated from the shrimp than the two crab species. For example, in an overlapping region of 490 bp, peJK1 and the shrimp MIH shared 47.6% homology while 44-47% sequence homology was observed in a smaller region of overlap (317-342 bp) of peJK1 with MIH from the other two crab species (*Carcinus maenas* and *Callinectes sapidus*). peJK1 also showed higher homology to the crustacean hyperglycemic hormone (CHH) nucleotide sequences of the crayfish, *O. limosus*, and the crab *C. maenas*, compared to MIH sequences from the two crab species. For comparisons with the CHH sequences, in a 400 bp overlap region, peJK1 shared 51.5% sequence homology with the crayfish CHH (section A2.4, Appendix 2), and 45.4% sequence homology in a 509 bp overlap with the crab CHH cDNA sequence (Table 3.2).

Similar results were also noted for cDNA clones peJK2 and peJK3. Both shared higher sequence homology to the MIH from the shrimp and CHH of both the crayfish and crab compared to the MIH from the two crabs. The sequence homology shared between peJK2 and peJK3 with the shrimp MIH, and CHH were quite similar, between 45 to 50% sequence identity, in a region of overlap > 379 bp (Table 3.2*). When compared to MIH from both crab species, the homology detected in ca. 280 bp overlap between the peJK2 and peJK3 with the MIH was between 48 to 51% sequence identity (Table 3.2). The other interesting features were that both the crab MIH sequences shared high sequence identities (82.7% over 324 bp overlap), whereas the MIH sequence isolated from the shrimp had higher sequence identity to CHH sequences from both the lobster and crayfish, compared to the MIH sequence from both crab species (Table 3.2).

Table 3.2. Sequence homology of cDNA clones peJK1, peJK2 and peJK3 compared to nucleotide sequences of other crustacean eyestalk neuropeptides. The cDNA clones were compared with the cDNA sequence encoding the putative MIH and signal peptide sequence.

Sequence	animal	sequence	animal	% homology	bp O/L
peJK1	lobster[a]	MIH	shrimp[b]	47.6%	490*
585 bp		MIH	blue crab[c]	46.8%	342
		MIH	shore crab[d]	44.2%	317
		CHH-A	crayfish[e]	51.5%	400*
		CHH	lobster[f]	48.8%	297
		CHH	shore crab[g]	45.4%	509*
peJK2	lobster[a]	peJK3	lobster[a]	96.6%	558
568 bp		MIH	shrimp	46.1%	436*
		MIH	blue crab	50.2%	287
		MIH	shore crab	51.4%	280
		CHH-A	crayfish	47.8%	542*
		CHH	lobster[f]	47.0%	349
		CHH	shore crab	45.4%	509*
peJK3	lobster[a]	MIH	shrimp	47.7%	379
568 bp		MIH	blue crab	48.6%	288
		MIH	shore crab	50.5%	279
		CHH-A	crayfish	49.8%	526*
		CHH	lobster[f]	49.0%	352
		CHH	shore crab	45.2%	586*
MIH	shrimp[b]	MIH	blue crab	50.7%	268
484 bp		MIH	shore crab	49.4%	324
		CHH-A	crayfish	53.8%	403
		CHH	lobster[f]	63.8%	265
MIH	shore crab[d]	MIH	blue crab	82.7%	324
342 bp		CHH-A	crayfish	50.9%	283
		GIH	lobster[h]	62.2%	344

[a] *Jasus edwardsii*; [b] *Penaeus vannamei* (Sun, 1994); [c] *Callinectes sapidus* (Lee *et al.*, 1995); [d] *Carcinus maenas* (Klein *et al.*, 1993b); [e] *Orconectes limosus* (de Kleijn *et al.*, 1994a); [f] *Homarus americanus* (de Kleijn *et al.*, 1995); [g] *C. maenas* (Weidemann *et al.*, 1989) [h] *Homarus americanus* (de Kleijn *et al.*, 1994). MIH (moult-inhibiting hormone), CHH (crustacean hyperglycemic hormone), GIH (gonad-inhibiting hormone). * sequence homology in a larger region of bp overlap.

The cDNA nucleotide sequences were also compared to the sequence of the 947 PCR fragment. All three cDNA sequences show between 46-48% homology when compared to the entire 947 PCR sequence, in both normal and complementary strand orientation. When the search was concentrated on the coding region, based on the predicted protein coding calculations (nucleotide 240-531), the homology ranged from 46-50% using both normal and complementary strand orientation in that region (Table 3.3). Similar results were also obtained for the predicted non-coding regions of the 947 PCR sequence, in both strand orientation. None of the searches showed any significant stretch of contiguous nucleotides which were homologous between the 947 PCR sequence and the cDNA clones.

Table 3.3. Sequence homology comparison of the 947 PCR product to the cDNA sequences of clones peJK1, peJK2 and peJK3.

947 sequence[1]	cDNA sequence	normal % homology	bp overlap	comp. % homology	bp overlap
1-960	peJK1	46.6%	573	48.5%	388
	peJK2	47.3%	446	45.7%	529
	peJK3	48.1%	428	47.4%	521
1-240	peJK1	45.8%	203	45.8%	216
	peJK2	49.5%	188	48.2%	199
	peJK3	49.2%	195	46.5%	202
240-531	peJK1	50.7%	276	48.9%	276
	peJK2	46.5%	243	50.6%	237
	peJK3	45.9%	257	48.5%	237
531-960	peJK1	45.8%	179	46.5%	228
	peJK2	47.9%	411	45.3%	375
	peJK3	48.1%	428	48.2%	438

[1] The 947 PCR product was compared in both normal and complementary (comp.) strand orientation to the cDNA clones. Nucleotides 240-531 comprise a protein coding region; 531-960, a non-coding region; and 1-240, a non-coding region based on algorithm calculations of Fickett (1982) on DNASIS.

3.4 DISCUSSION

3.4.1 Isolation of the putative MIH gene sequence from the lobster genomic library

Homology searches conducted on the sequenced regions of plasmids p2A and p66 failed to locate either the 947 PCR sequence or any of the cDNA sequences of peJK1, peJK2 and peJK3 (Appendix 2). However, strong hybridization signals were shown when the restriction digests of these genomic clones were hybridized to the 947 PCR sequence.

Southern blot analysis of restriction digests of plasmid p66 showed that there were two *Acc*I and *Xba*I restriction fragments which were detected by the 947 PCR sequence. However, both these restriction sites were absent from the 947 PCR sequence. This suggests the presence of more than one 947 PCR related sequence within the genomic clone p66. This observation was supported in part, by Southern blot analysis of restriction digests of lobster DNA (section 2.3.4, Chapter II) where three *Bgl*II restriction fragments (6.5, 4.8 and 3.6 kb) were detected by the 947 PCR sequence. The 947 PCR sequence however, has only one internal *Bgl*II restriction site.

Based on similarity searches conducted using the BLAST tool, similarity to the *C. maenas* CHH precursor was detected. When examined in detail, both p2A and p66 contained CT rich regions which were also present in the *C. maenas* CHH precursor cDNA sequence. In a region of 1066 bp overlap, the 2A-4 region of plasmid p2A showed 48% sequence homology with the cDNA sequence of the *Carcinus maenas* CHH precursor (Weidemann *et al.*, 1989; accession no. emb:X17596). The 70 bp CT rich region extending from nucleotides 729 to 798 in 2A-4, is very similar to the CT rich region found in the *C. maenas* CHH precursor cDNA sequence (98% sequence identity over 100 bp overlap) (Appendix 2). In the CHH cDNA sequence, the 123 bp CT rich region extends from nucleotides 750 to 872, 236 bp downstream from the end of the sequence coding for the CHH mature peptide. A similar CT rich region, 54 bp long, was also found in the 66-2 region of p66, extending from nucleotides

1481-1543 (Appendix 2). This may be evidence to suggest that the genomic clones may contain a CHH like gene, as the CT rich region found in the *C. maenas* CHH percursor was derived from the coding sequence.

3.4.2 Sequence analysis of the cDNA clones

The most obvious feature was the 96.6% sequence homology shared between cDNA clones peJK2 and peJK3. The two sequences may well represent allelic forms of the same gene, rather than products of alternative splicing of a common mRNA precursor; several of the nucleotide differences between the two sequences involved the substitution of one base for another, and the insertion/deletion of several nucleotide bases. The CHH has been shown to be polymorphic, generally found as two closely related isoforms. cDNA clones that encode for two structurally different CHH neuropeptides have been isolated from *H. americanus* (92% identity) (Tensen *et al.*, 1991b; de Klein *et al.*, 1995), and *O. limosus* (de Kleijn *et al.*, 1994a).

Translation of the cDNA clones peJK2 and peJk3 also showed a high proportion of hydrophobic residues, notably valine and alanine, which occur as stretches of up to 8 consecutive valine residues in the deduced amino acid sequence for peJK3. The set of 6 and 8 consecutive valine residues near the N-terminus in peJK2 and peJK3 respectively, probably form the hydrophobic core of the signal peptide sequence. A similar set of 6 consecutive valine residue also reside in the signal peptide sequence of the *O. limosus* CHH cDNA sequences (de Kleijn *et al.*, 1994a). An acceptable cleavage site was assigned after amino acid 31 in peJK2, and 34 in peJK3, based on the (-3,-1)-rule, and comparisons to amino acid counts for eukaryotic signal counts (von Heijne, 1986). Based on this assignment, the extra amino acids encoded by peJK3 are found in the signal peptide, which is cleaved from the mature peptide once export is under way. The signal peptide itself is a transient N-terminal signal sequence found on most secretory proteins, and serves to initiate export across the endoplasmic reticulum (von Heijne 1986; Habener 1987; Lehninger *et al.*, 1993).

The hydrophobic profile of the deduced amino acid sequences for peJK2 and peJK3 shows that the N-terminal and central domain of these putative peptides are hydrophobic. This feature may be important in the secondary and tertiary structure of the protein, as well as determining the environment in which the protein is found. The hydrophobic domain could be a membrane-buried segment or buried within the protein itself (von Heijne, 1987; Taylor, 1987; Lehninger *et al.*, 1993). The hydrophobicity plot for peJK1 was not as distinctive as the other two.

When the three cDNA sequences were compared to available cDNA sequences of other crustacean eyestalk neuropeptides, peJK1, peJK2 and peJK3 showed higher sequence identity to the MIH-like neuropeptide of the shrimp, and the CHH of the crayfish and lobster, in contrast to the putative MIH sequence from the two crabs. The differences seen in MIH sequence identity between the two crabs (82%), and the shrimp compared to the crabs (49%), could be attributed to the probes used in the isolation of the shrimp putative MIH sequence. Chang *et al.* (1990) isolated a hydrophobic peptide, 71 amino acid residues long, with both moult-inhibiting hormone activity and significant hyperglycemic activity from the lobster *H. americanus*. The MIH-like neuropeptide from the shrimp *P. vannanmei* was isolated by screening a cDNA library constructed by 3′ and 5′ rapid amplification of cDNA ends (RACE) method, using primers derived from the conserved region of the MIH peptide of *H. americanus*. The biological activity of this encoded peptide has not been tested (Sun, 1994).

The CHH and MIH neuropeptides are considered to be related (reviewed in Keller, 1992), and therefore the higher sequence homology of peJK1, peJK2 and peJK3, to the sequences encoding the shrimp MIH-like neuropeptide and the CHH peptides is not unexpected. However, the deduced amino acid sequences encoded by peJK1, peJK2, and peJK3 do not show the same pattern of conserved amino acid sequences when compared with those of the CHH/MIH/VIH peptide family. In this family, the six invariant cysteine residues are located in the same positions, and conserved amino acid residues are located in the vicinity of the six cysteine residues (Sun, 1994; de Kleijn and Van Herp, 1995 (review)). There are, however, differences

in the deduced amino acid sequences of the putative MIH peptides isolated from the four different species. The *P. vannamei* deduced MIH sequence has 49% and 29% homology to the *H. americanus* MIH, and *C. maenas* MIH respectively, whereas both crabs, *Callinectes sapidus* and *C. maenas*, share 78% sequence homology.

The dissimilarity seen between the deduced amino acid sequences of peJK1, peJK2 and peJK3, and the published putative MIH peptides may well reflect interspecific differences. However, until the biological activity of the peptides encoded by peJK1, peJK2 and peJK3 are examined, it is not possible to ascertain whether the cDNA clones isolated actually encode putative MIH gene sequences. Preliminary characterization of the proteins encoded by the cDNA clones was not carried out. By sequencing, all of the cDNA clones isolated were found to have the inserts cloned in the wrong orientation to the transcription promoters present in the plasmid vector, and hence, expression of the encoded proteins in a coupled transcription translation cell free system could not be performed. The cDNA clones, however, have been characterized in terms of their expression in different tissues of the lobster (discussed in Chapter IV).

CHAPTER IV

Expression of the cDNA clones peJK1, peJK2 and peJK3, in the various tissues of the lobster

4.1 INTRODUCTION

Based on the immunological evidence provided by Mattson and Spaziani (1985a) that the crustacean MIH is related to the vasopressins, PCR was used to isolate the putative MIH gene sequence from lobster DNA, using primers derived from conserved regions of the rat vasopressin gene (Chapter II). Three cDNA clones, peJK1, peJK2 and peJK3, were isolated from a lobster eyestalk cDNA library using the 947 PCR fragment as a probe. Sequence analysis showed that these cDNA sequences had 44-51% sequence identity to published cDNA sequences of the putative MIH, and CHH cDNA sequences.

The putative MIH and CHH are produced in the X-organs sinus gland system of the eyestalk, as demonstrated by immunocytochemical studies, and *in situ* hybridization studies (Dircksen *et al.*, 1988; Kallen and Meusy, 1989; de Kleijn *et al.*, 1992; Klein *et al.*, 1993a; 1993b). In this chapter tissue specific expression of the three cDNA clones were examined by northern blot analysis, and *in situ* hybridization studies, to determine the distribution of the cloned cDNA sequences and whether the localization of these sequences were similar to the predicted sites of putative MIH synthesis.

4.2. MATERIALS AND METHODS

4.2.1 Tissue specific expression of the cDNA clones by northern analysis

4.2.1.1 Preparation of cDNA probes

cDNA inserts were excised from the pSPORT plasmid, by overnight digestion at 37°C with *Not*I restriction endonuclease (3.5 U/µg). The cDNA inserts were gel purified, and DNA isolated by electroelution using the BIO-TRAP apparatus (Schleicher and Schule). Twenty five nanograms of cDNA insert was labelled with [α-^{32}P] dCTP (50 µCi, 3000 Ci/mmol), to 10^7 to 10^8 cpm/µg using the random prime labelling kit (Boehringer Mannheim). The radiolabelled probe was purified by column chromatography before use.

4.2.1.2 Northern hybridization of mRNA and total RNA to the cDNA clones

Northern blotting of total RNA and poly(A)$^+$ isolated from the epithelia, gill, heart, hepatopancreas and abdominal flexor muscle of an individual lobster, and from lobster eyestalks was described in section 2.2.3.3 (Chapter II).

Northern hybridization of the isolated poly(A)$^+$ was carried out according to Lee *et al.* (1992) but with a final wash at a higher stringency of 0.1 x SSC/0.1% SDS at 65°C for 15 min. The membrane was exposed to X-ray film at –80°C with intensifying screens for 2 h, 4 h, and 14 h. The northern blot was probed with each of the three cDNA inserts. After each hybridization, the membrane was stripped of the hybridized probe by washing in 0.1 x SSC/0.1% SDS at 95°C for 2 x 20 min, then in 0.1% SDS at 100°C till the solution cooled to room temperature. The membrane was checked by overnight exposure to X-ray film.

4.2.2 Tissue specific expression of cDNA sequences using *in situ* hybridization to frozen sections

A lobster was chilled on ice for 40 min, and the following tissue was dissected: epithelial tissue above the dorsal heart, eyestalk, gill, heart, and abdominal flexor muscle in the cephalothorax region. The different tissues were fixed in 4% paraformaldehyde at 4°C for 50-60 min, washed twice in phosphated-buffered saline

PBS, and left overnight in 15% sucrose in PBS (described in detail in section 2.2.3.2, Chapter II). The tissue blocks were cut in the following planes: epithelia, transverse and longitudinal sections; eyestalk, longitudinal section; gill, heart and muscle, transverse sections. Cryostat cut sections (16 µM thick, –26°C) of the different tissues were stored at –80°C for 2-3 months.

In situ hybridization of the tissue sections to the cDNA probes were performed as described in section 2.2.3.2 (Chapter II). [α-^{32}P] dCTP labelled cDNA probes were prepared as described in section 4.2.1.1, to a specific activity of 10^8 cpm/µg DNA.

Controls (RNase digested sections) for each tissue were included in these sets of experiments. The activity of RNase A (100 µg/ml) was first tested by its ability to degrade ten micrograms of total RNA at 37°C for 1 h, and the results visualized by agarose gel electrophoresis. Tissue sections on control slides were covered with 500 µl of 2 x SSC containing 100 µg/ml RNase A, and incubated in a moist environment at 37°C for 1 h. The control tissue sections were then prehybridized and hybridized as described for the experimental sections. Contamination of experimental slides with RNase was avoided by keeping the two sets of slides separate during the course of the experiment.

Following post-hybridization washes, tissue sections were exposed to X-ray film for 45 min, 2.5 h and 16 h, to determine the length of exposure to liquid autoradiographic emulsion. Both control and experimental slides were developed after a 3 day exposure period (29 x exposure to the X-ray film).

4.3 RESULTS

4.3.1 Distribution of peJK1, peJK2 and peJK3 mRNAs

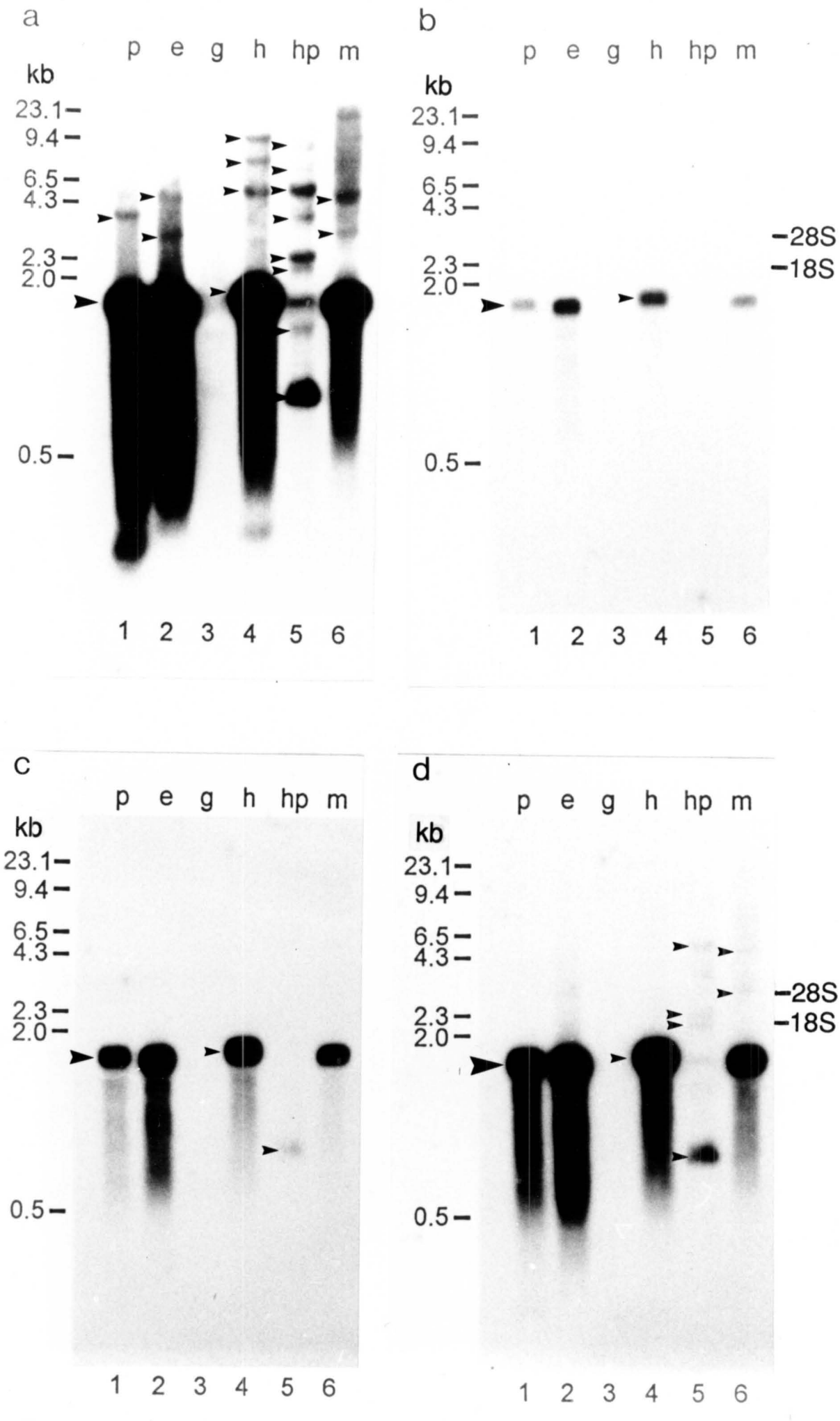
Northern blot analysis results presented here were based on hybridization of the cDNA inserts to poly(A)$^+$ RNA. The amount of poly(A)$^+$ RNA loaded varied for the different tissues: epithelia, 1.0 µg; eyestalk, 1.44 µg; gill, 2 µg; heart, 1.35 µg; hepatopancreas, 2.7 µg; and muscle, 2.43 µg. Varying amounts of poly(A)$^+$ RNA were loaded as the concentration of the poly(A)$^+$ RNA solutions were not measured, but calculated based on spectrophotometry readings at A_{260} of the original solution eluted from the separation system. These solutions could have been contaminated with a trace of paramagnetic particles used in the purification of mRNA (section 2.2.3.3 (ii), Chapter II). Subsequently, the concentrations of poly(A)$^+$ RNA solutions were determined both by spectrophotometry readings at A_{260}, and by comparison with tRNA standards spotted onto the surface of a 1% agarose gel containing 0.5 µg/ml ethidium bromide, and visualized over a UV transilluminator (Sambrook *et al.*, 1989).

To compensate for the varying amounts of poly(A)$^+$ RNA loaded for each tissue, hybridization bands on the autoradiogram were analysed semi-quantitatively by measuring the optical densitometry reading of each band with an analytical scanner (AAB, Advanced American Biotechnology). The densitometry readings were affected both by the scanning sensitivity, and delineating the perimeter of the bands. Because of this, the autoradiogram was scanned at three sensitivity setting, and densitometry readings for that particular scan was normalized for 1 µg of poly(A)$^+$ RNA and expressed as a percentage of the eyestalk band in that particular scan.

4.3.1.1 Northern hybridization to peJK1

Northern hybridization to peJK1, using the high stringency wash of 0.2 x SSC/0.1% SDS at 60°C (Lee *et al.*, 1992), gave a heavy band at approximately 1.8 kb extending as a smear past 0.5 kb for epithelial, eyestalk, heart and muscle poly(A)$^+$ RNA, after 22 h exposure (Fig. 4.1a). There were also larger mRNAs hybridizing to peJK1; 3.9 kb transcript for epithelial tissue, 5.3 and 2.9 kb for the

Figure 4.1. Northern blot analysis of poly(A)$^+$ RNA from epithelial tissue (p), eyestalk (e), gill (g), heart (h), hepatopancreas (hp), and muscle (m). The blot was hybridized with the cDNA probe peJK1; wash stringency of 0.2 x SSC/0.1% SDS at 60°C, and exposed for 22 h (**a**). The wash stringency was increased to 0.1 x SSC/0.1% SDS at 65°C, and the blot exposed for 2 h (**b**), 4 h (**c**) and 13.5 h (**d**). The positions of the 28S and 18S rRNA bands are indicated on the right, and single stranded *Hin*dIII digested λ DNA size markers are indicated on the left (kb): 23.1, 23,130 bp; 9.4, 9,416 bp; 6.5, 6,557 bp; 4.3, 4,316 bp; 2.3, 2,322 bp; 2.0, 2,027 bp; 0.5, 564 bp. Arrows locate the position of the bands; big arrows indicate the 1.7 kb band found in the epithelia, eyestalk and muscle extracts.


a
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
1 2 3 4 5 6
b
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
28S
18S
1 2 3 4 5 6
c
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
1 2 3 4 5 6
d
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
28S
18S
1 2 3 4 5 6

eyestalk, 11.2, 7.8 and 5.3 kb for heart tissue, and 4.8 and 3.1 kb for the muscle tissue. PeJK1 hybridized to 9 major bands (9.0, 7.5, 6.0, 3.6, 2.4, 2.2, 1.7, 1.4, 0.9 kb) in the hepatopancreas poly(A)$^+$ RNA.

When the wash stringency was increased to 0.1 x SSC/0.1% SDS at 65°C, hybridization of the probe to the larger bands was reduced. After 2 h exposure, a single band at 1.7 kb was seen for epithelial, eyestalk and muscle tissue, and 1.8 kb for heart (Fig. 4.1b). Although the single bands in Fig. 4.1b are slightly different in their mobilities, the bands seen for the epithelia, eyestalk, and muscle were standardised to 1.7 kb size, and the heart, 1.8 kb. The small differences may be due to the gel run rather than an actual size difference in the mRNA transcripts detected. The size of mRNA transcripts were estimated by comparing their relative mobilities to single stranded DNA size markers run on the same gel.

The bands seen in the heart and eyestalk poly(A)$^+$ RNA were the heaviest, indicating a higher level of peJK1 mRNA transcripts were present in these tissues (Fig. 4.1b). The relative intensities of the bands as compared to the eyestalk bands were: heart, 104%; epithelia, 51%; and muscle, 24% of the eyestalk band (Table 4.1)

Table 4.1. Relative intensities of northern hybridization bands detected by peJK1 (Fig. 4.1b). Optical densitometry measurements of the bands were normalized to 1 µg, and expressed as a percentage of the eyestalk band.

Tissue	Band size (kb)	Mean % ± s.d.[a]
epithelia	1.7	50.9 ± 8.4
eyestalk	1.7	100
heart	1.8	104.7 ± 1.1
muscle	1.7	24.5 ± 3.1

[a] The standard deviation values reflect the amount of variation obtained from this method. The autoradiogram was scanned at three different sensitivity settings on the scanner. Densitometry measurements were then determined from the scanned images several times (n=6), to reduce the variation due to the manual marking of the bands.

After 4 h exposure, a single band (0.90 kb) was seen in the hepatopancreas, and the other larger ones were observed after 14 h exposure (Fig. 4.1c and d). The smear seen extending from approximately 1.8 kb to 0.5 kb in the epithelial, eyestalk, heart and muscle tissue indicates the presence of mRNA transcripts in that size range.

4.3.1.2 Northern hybridization to peJK2 and peJK3

Northern blot analysis using cDNA clones peJK2 and peJK3 as probes, revealed specific hybridization to a single band of poly(A)$^+$ RNA (0.68 kb) in the eyestalk after 2 h and 4 h exposure (Fig. 4.2a and b). Similar results were obtained using peJK3 (results not shown). A faint band (0.70 kb) was seen in the epithelial tissue lane after 4 h (Fig. 4.2b). After 14 h, the intensity of the 0.7 kb epithelial band was approximately 3 fold less than the major eyestalk band, and 2 fold and 5 fold greater than the heart, and the hepatopancreas and muscle bands respectively (Table 4.2).

Both the eyestalk major band, 0.68 kb and the epithelial band, 0.70 kb were standardized to 0.70 kb; the slight variation in band sizes were very similar to that observed with peJK1 and may well reflect the gel electrophoresis conditions.

After 14 h exposure, more bands were detected in the epithelial tissue (1.7, 0.40, 0.30 kb) and eyestalk (1.7, 0.43 kb). PeJK2 and peJK3 also hybridized to at least 3 bands (2.16, 1.8, and 0.34 kb) in the heart, 4 bands (6.0, 2.4, 2.2 and 0.90 kb) in the hepatopancreas, and at least 4 bands (5.35, 1.7, 0.70 and 0.36 kb) in the muscle (Fig. 4.2c and d). Some of these bands correspond to the major bands (1.7-1.8 kb, 0.9 kb) detected by the peJK1 probe. The average intensities of the bands detected by peJK2 and peJK3, expressed relative to the major eyestalk band, showed that peJK2 and peJK3 mRNA transcripts were produced primarily in the lobster eyestalk (Table 4.2).

Figure 4.2. Northern blot analysis of poly(A)$^+$ RNA from epithelial tissue (p), eyestalk (e), gill (g), heart (h), hepatopancreas (hp), and muscle (m). The blot was hybridized to cDNA probe peJK2, and a final high stringency wash of 0.1 x SSC/0.1% SDS at 65°C was included. The blot was exposed for 2 h (**a**), 4 h (**b**) and 14 h (**c**). After stripping the membrane, the blot was hybridized to cDNA probe peJK3 using the same conditions and exposed for 14 h (**d**). The positions of the 28S and 18S rRNA bands are indicated on the right, and single stranded *Hin*dIII digested λ DNA size markers are indicated on the left. The positions of the bands are indicated by arrows; big arrows indicate the major eyestalk band (0.7 kb). Bars indicate the position of faint bands seen on the autoradiogram.

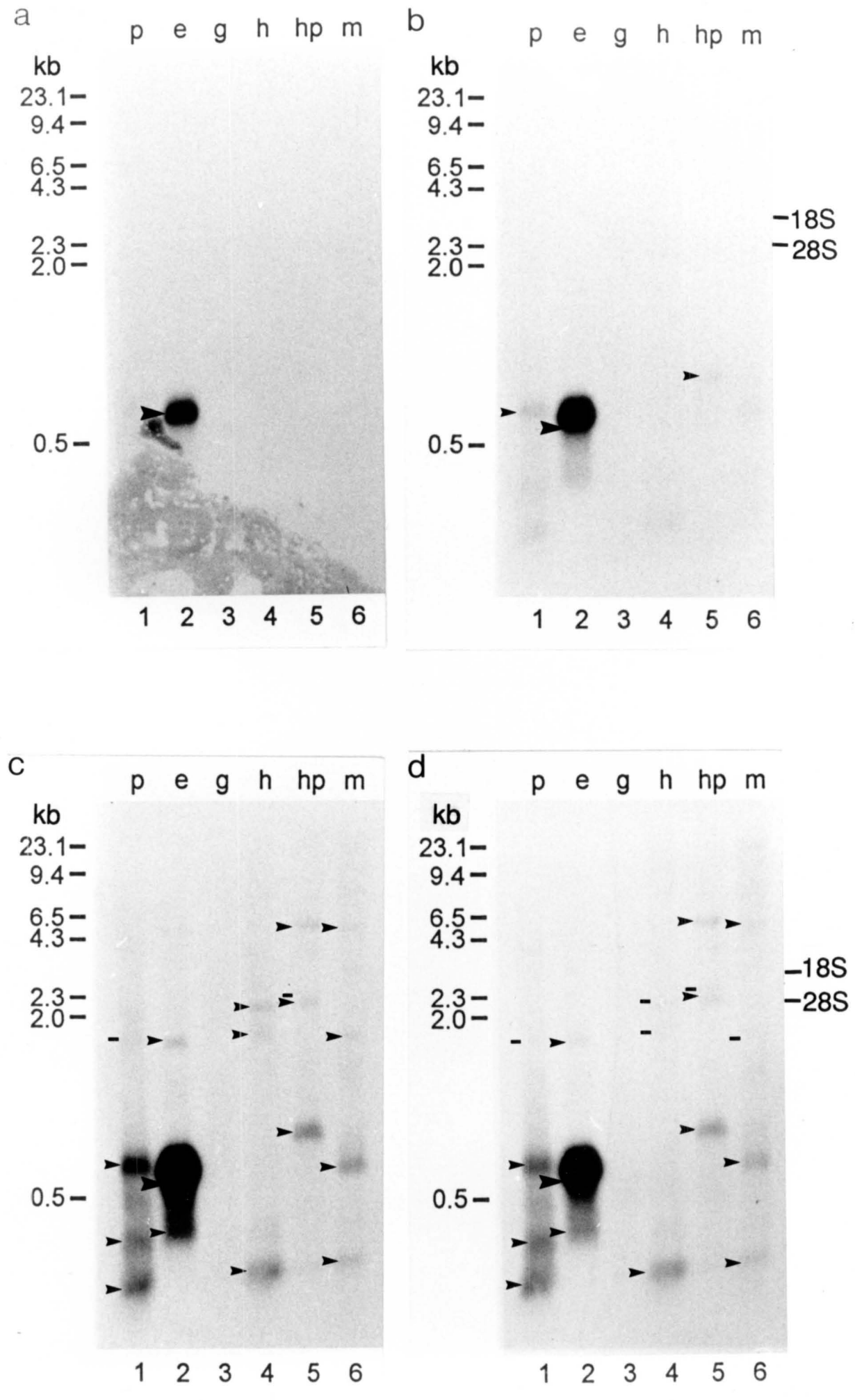

a
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
1 2 3 4 5 6
b
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
18S
28S
1 2 3 4 5 6
c
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
1 2 3 4 5 6
d
p e g h hp m
kb
23.1
9.4
6.5
4.3
2.3
2.0
0.5
18S
28S
1 2 3 4 5 6

Table 4.2. Relative intensities of northern hybridization bands detected by peJK2 and peJK3, at 13.5 h exposure (Fig. 4.2c and d). Optical densitometry measurements of the more obvious bands were normalized to 1 µg, and expressed as a percentage of the major eyestalk band.

Tissue	Band size (kb)	Mean % ± s.d.[a]
epithelia	1.7	3.1 ± 0.6[b]
	0.7	29.0 ± 3.0
eyestalk	1.7	3.0 ± 0.5[c]
	0.7	100
heart	0.34	14.8 ± 2.3
hepatopancreas	6.0	2.0 ± 0.6
	0.9	4.9 ± 0.6
muscle	0.7	4.7 ± 0.9

[a] The standard deviation reflects the amount of variation obtained when both autoradiograms were each scanned using three different sensitivity settings. Unless indicated, n=6.

[b] Mean % based on three readings.

[c] Mean % based on four readings.

4.3.1.3 Comparison of the mRNA transcripts detected using the different probes

The northern blot membrane was probed successively with the three cDNA clones as well as the 947 PCR fragment. Although the sizes of the detected mRNA transcripts were only estimates, the relative positions of the bands detected by the cDNA clones and the PCR product could still be compared as the same northern blot was used for all four experiments. The sizes of the mRNA transcripts detected are summarized in Table 4.3. Most of the bands detected using the cDNA clones, also hybridized to the 947 PCR fragment. The more abundant mRNA transcripts detected by peJK1 in the epithelia, eyestalk, heart, hepatopancreas and muscle, also hybridized to both peJK2 and peJK3. The intensities of the peJK1 related bands detected by peJK2 and peJK3, however, were much less than the major eyestalk band (Table 4.2).

Table 4.3. Comparison of mRNA transcripts detected using the following probes: cDNA clones peJK1, peJK2 and peJK3; and the 947 PCR product. The more abundant transcripts are underlined.

Tissue	size of mRNA transcripts (kb) detected using:[a]		
	peJK1	peJK2, peJK3	947 bp product
epithelia	1.7	1.7, 0.7, 0.40, 0.30	1.7
eyestalk	1.7	1.7, 0.7, 0.43	1.7, 0.7
heart	1.8	2.16, 1.8, 0.34	1.8, 0.34
hepatopancreas	6.0, 3.6, 2.4, 2.2, 1.75, 0.90	6.0, 2.4, 2.2, 0.90	6.0, 2.2, 0.90
muscle	5.35, 3.10, 1.7,	5.35, 1.7, 0.7, 0.36	5.35, 1.7, 0.36

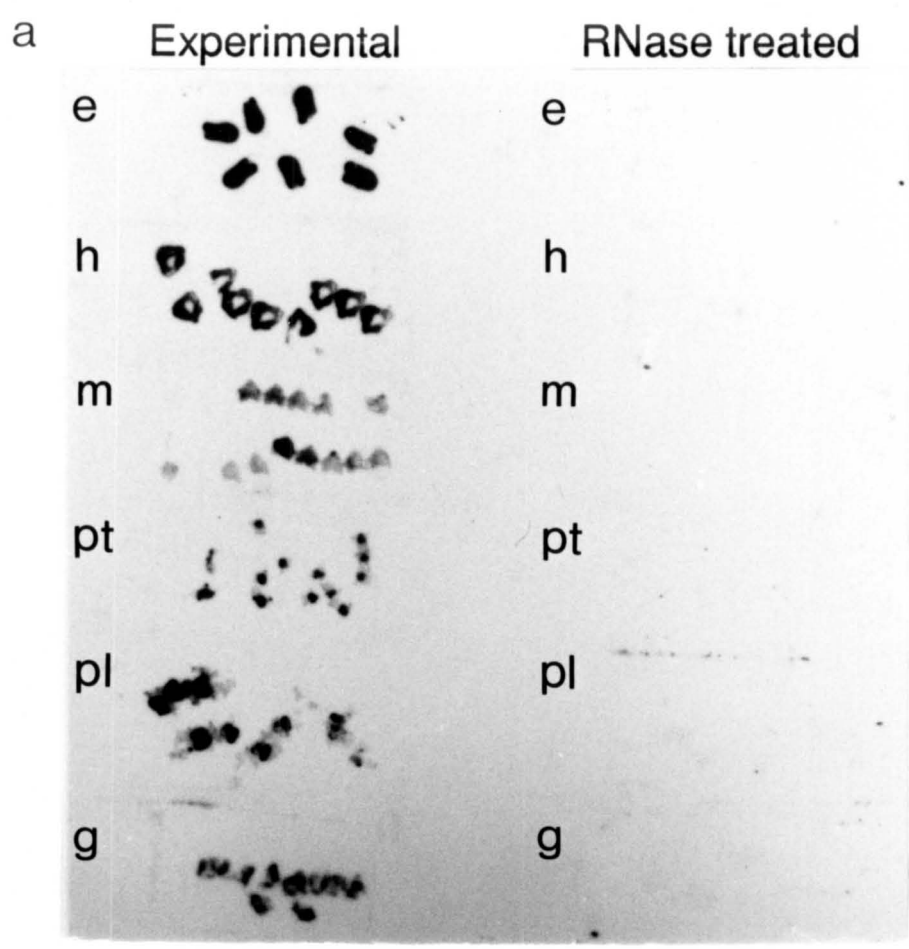
[a] Band sizes were standardised as explained in the text previously.

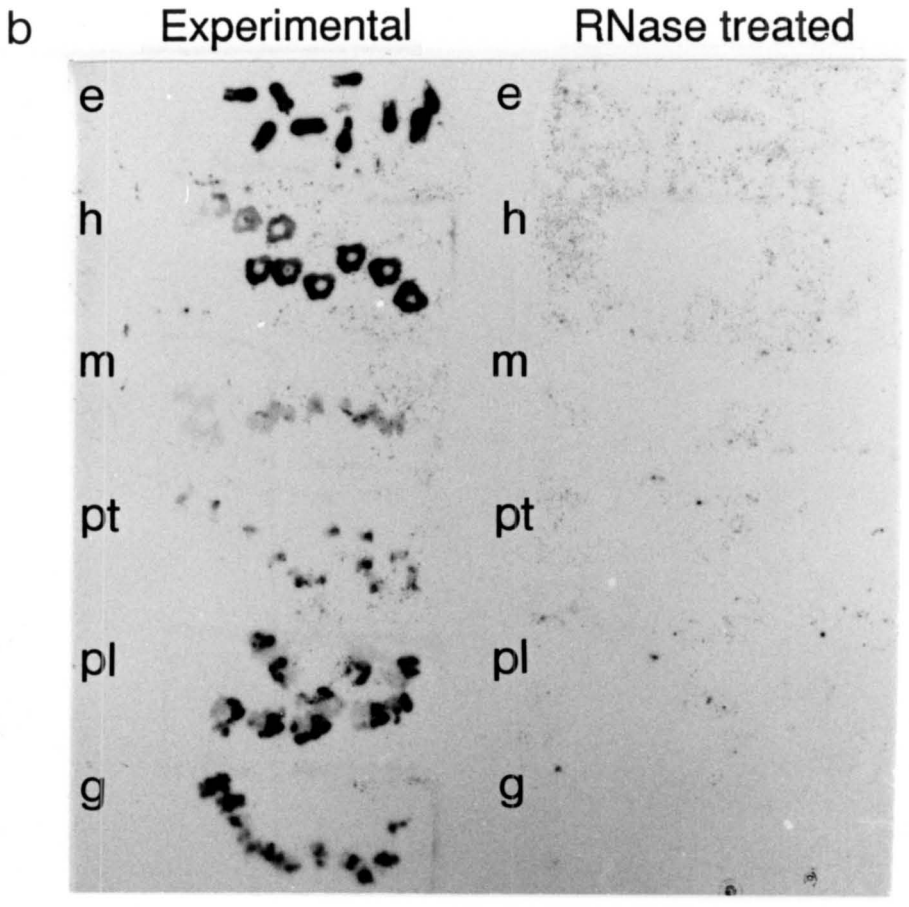
4.3.2 *In situ* hybridization using cDNA clones peJK1 and peJK2

The expression of the isolated cDNA sequences detected in the tissues in northern blot analysis, was further examined by *in situ* hybridization to determine the exact site of expression of the cDNA detected mRNA transcripts. The localization of the probes to specific regions within each tissue was evident on X-ray film after 2.5 h exposure (Fig. 4.3a and b). RNase digestion of the tissues abolished probe binding to background levels. The hybridization signal was strongest in the eyestalk. The signal obtained with peJK1 was much stronger than peJK2, even though both experiments were done at the same time and the probes were labelled to similar specific activities. RNase digestion of tissue sections reduced probe hybridization signals to background levels on X-ray film, after 16 h exposure (figures not shown).

4.3.2.1 Structure of the lobster eyestalk

Based on comparisons to the literature (Bliss and Welsh, 1952; Bullock and Horridge, 1965; Smith and Naylor, 1972; Rudolph and Spaziani, 1990) the following structures of the eyestalk were identified (Fig. 4.3). In the longitudinal sections of the lobster eyestalk, the 4 optic neuropiles, structures which relay information from the retina to the brain, were distinguishable; the lamina ganglionaris (LG), medulla


a
Experimental
RNase treated
e
h
m
pt
pl
g
e
h
m
pt
pl
g


b
Experimental
RNase treated
e
h
m
pt
pl
g
e
h
m
pt
pl
g

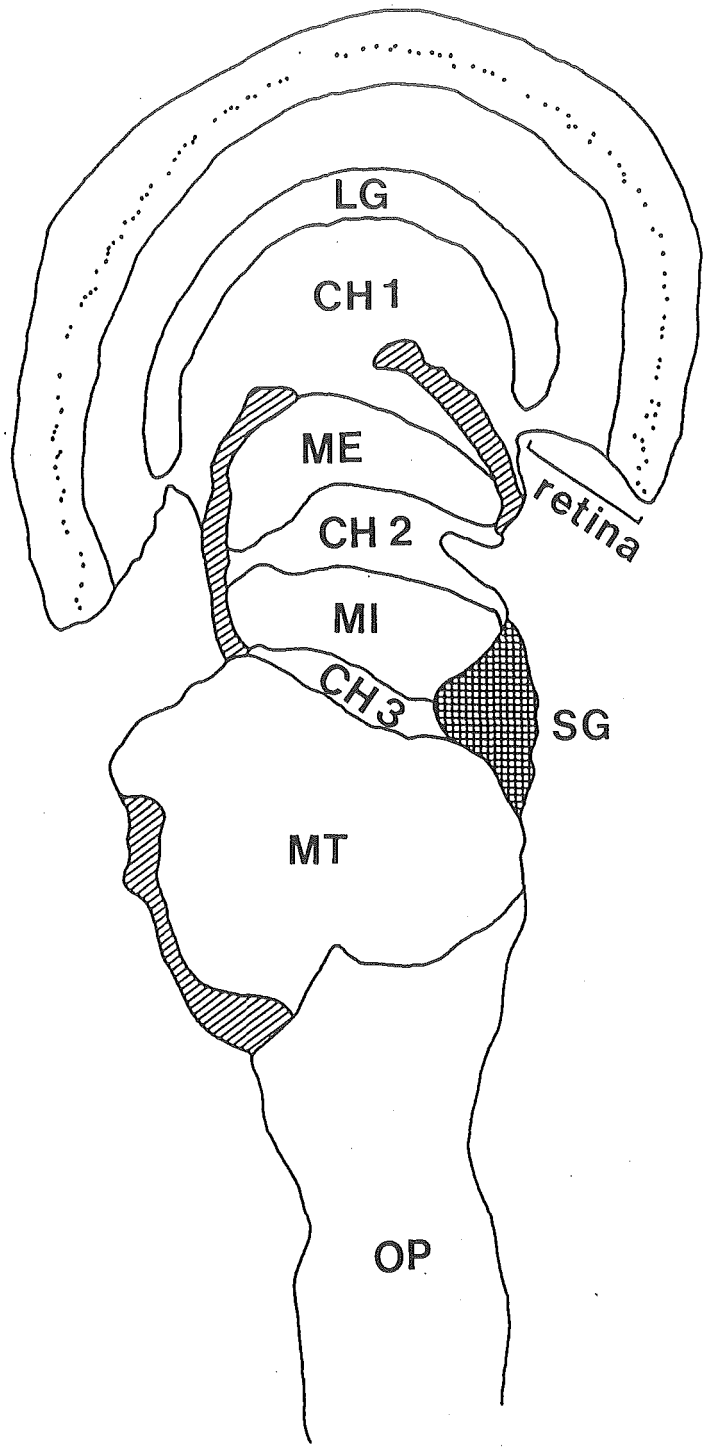


Figure 4.4. Diagram illustrating the various structures identified in the longitudinal section of the eyestalk of the lobster *J. edwardsii*. LG, lamina ganglionaris; ME, medulla externa; MI, medulla interna; MT, medulla terminalis; CH 1, chiasma 1; CH 2, chiasma 2; CH 3, chiasma 3; SG, sinus gland; OP, optic lobe peduncle. The neurosecretory regions in the eyestalk, and the sinus gland are indicated (shaded areas).

externa (ME), medulla interna (MI), and the medulla terminalis (MT). The optic neuropiles are connected to each other by fibres which form the first, second and third chiasma (CH1, CH2, and CH3); CH1 is between the lamina ganglionaris and medulla externa, and so on (Bullock and Horridge, 1965). The sinus gland was identified by the appearance and relative positions compared to that found in the published literature, and not by histochemical techniques.

Based on the classification of the different neurosecretory cell types in the eyestalk of the decapod *Carcinus maenas* (Smith and Naylor, 1972), several different cell types were distinguishable on the lobster eyestalk sections. Cell types 1 and 2, which were the smaller and densely packed neurosecretory cells, were abundant in the neurosecretory cell groups of the 3 neuropiles (ME, MI, MT). The size of the nucleus was large relative to the perikaryon (cell body). Interspersed amongst these cells were the larger cells (cell types 3-6), which have large nuclei and perikaryon; the cell boundaries are distinguishable in these cells.

4.3.2.2 Distribution of peJK1 in the lobster eyestalk

PeJK1 hybridized to mRNA within specific regions of the eyestalk, as demonstrated by the dense accumulation of silver grains where the probe had hybridized (Fig. 4.5a whole mount). RNase treatment of the control sections abolished probe hybridization, thus indicating that the probes were hybridizing to mRNA on the experimental tissue sections (Fig. 4.5c, d, h, i, j).

Examination of the sections in detail revealed that peJK1 was expressed in the neurosecretory cell groups of the medulla externa, medulla interna, and the medulla terminalis X-organs (MTXO) (located in the proximal ventral portion of the medulla terminalis) (Fig. 4.5b, f, and g). The neurosecretory cells of the medulla terminalis (Fig. 4.5g) have been referred to as the pars ganglionaris X-organs by Carlisle and Passano (1953), medulla terminalis X-organs by Knowles and Carlisle (1956) and the organ of Hänstrom by Gabe (1966) (cited in Smith and Naylor, 1972).

Figure 4.5. *In situ* hybridization of lobster eyestalk (lobster 1) to cDNA probe peJK1. Comparison of peJK1 probe hybridization on experimental slide (**a, b, e, f, g**), to corresponding regions on the RNase treated control sections (**c, d, h, i, j**). Probe hybridization is evident as accumulation of dense silver grains over the neurosecretory regions of the eyestalk in the whole mount (**a**). In the neurosecretory regions of the medulla externa and medulla interna (**b, f**), as well as the medulla terminalis (**g**), probe localization occurs on the nuclei of some of the cells, and only in the cytoplasm of the other neurosecretory cells. In the neurosecretory cells above the lamina ganglionaris, and retinal cells, peJK1 hybridized to the cell nuclei, and to the cytoplasm of the ommatidial cells (**e**). The corresponding region on the RNase-treated control is shown (**h**). The sections have been differentiated by staining with Haemotoxylin; cell nuclei appear as purple to pink structures, whereas cell boundaries are not as distinctive in the sections. *ME* medulla externa; *MI* medulla interna; *MT* medulla terminalis; *LG* lamina ganglionaris; *OM* ommatidia. *Bars*: a, c = 500 µm; b, d, e, f, h, and i = 100 µm; g = 50 µm.

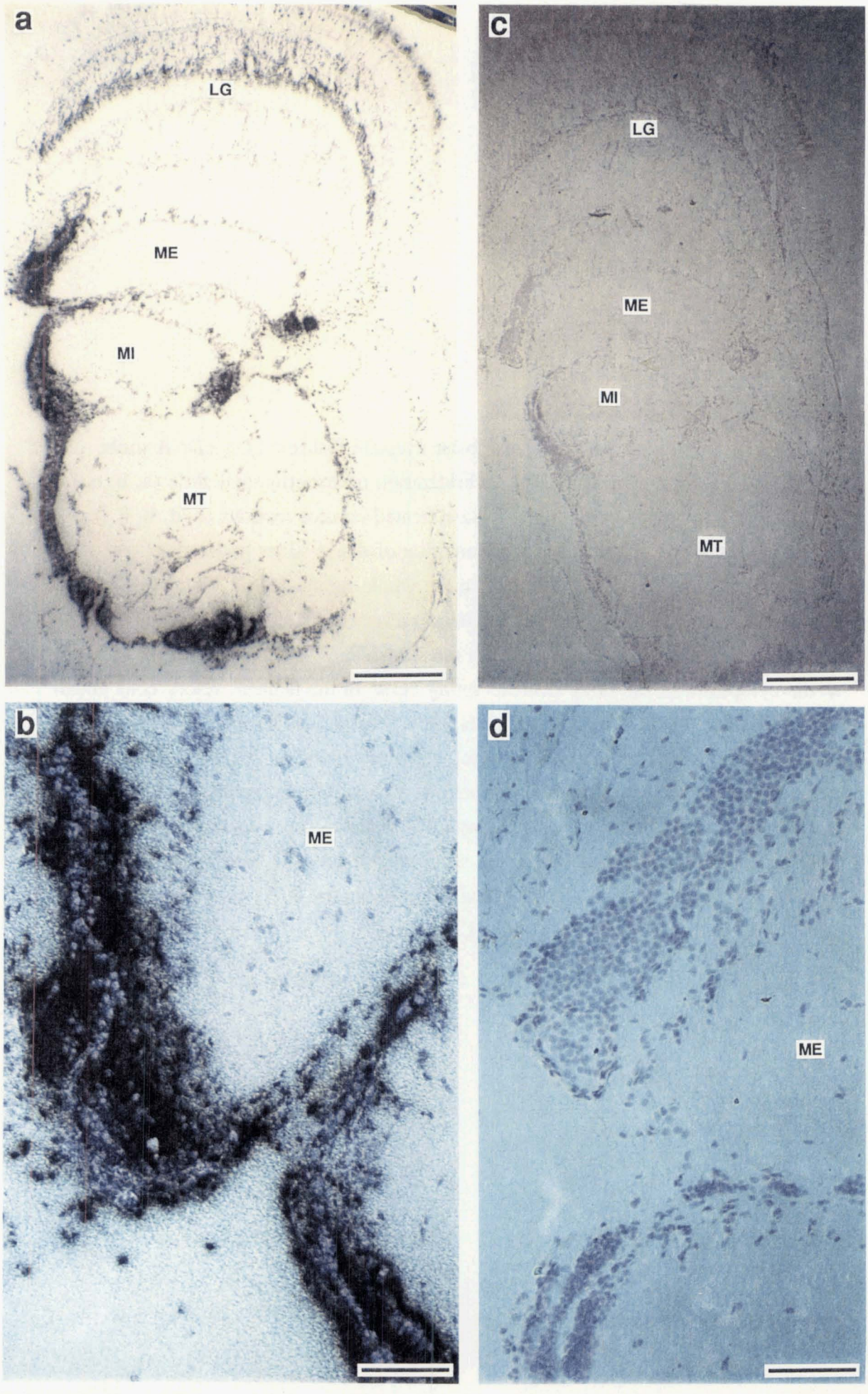
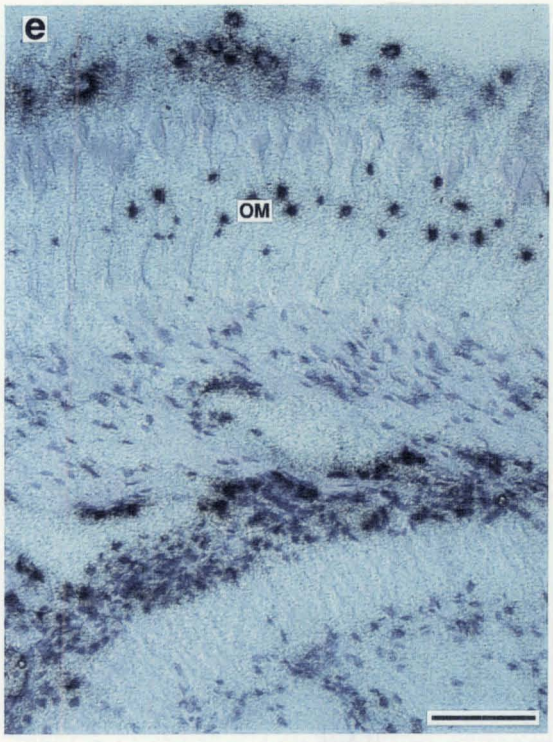
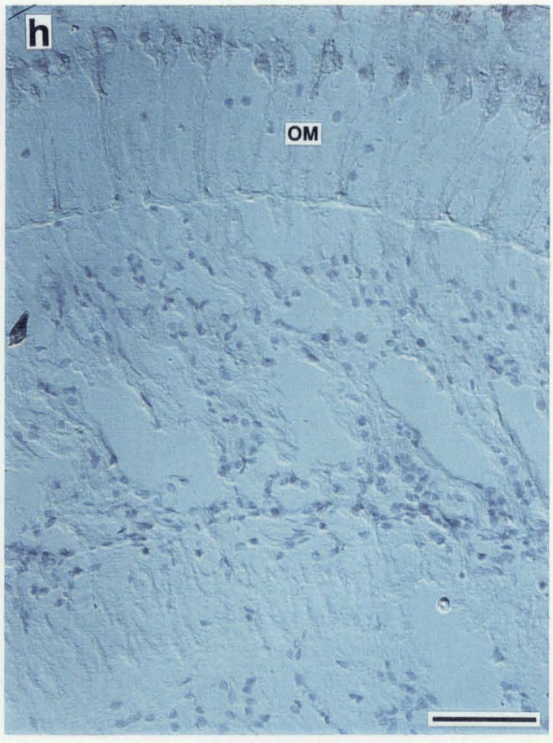
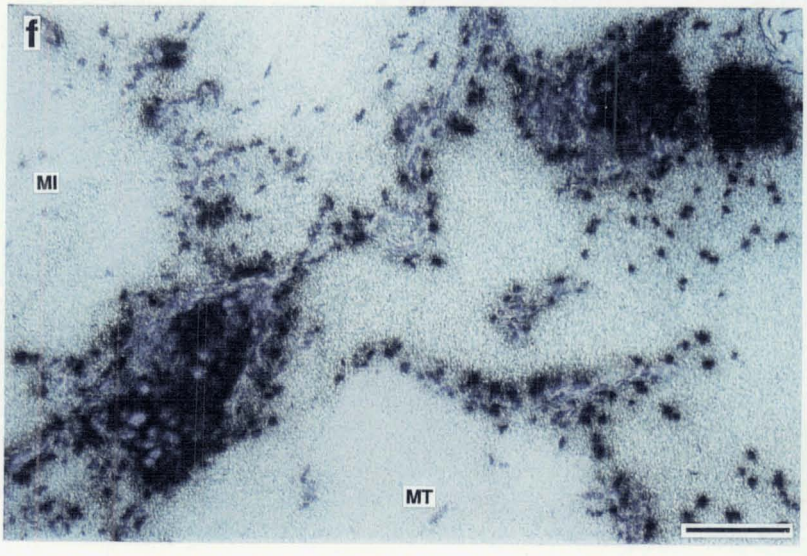
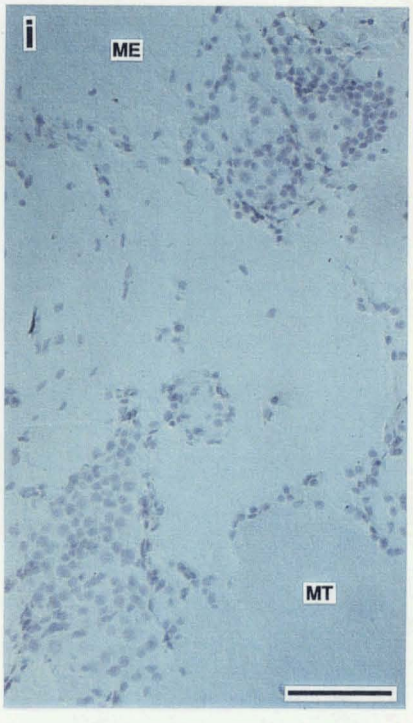

a
LG
ME
MI
MT
c
LG
ME
MI
MT
b
ME
d
ME

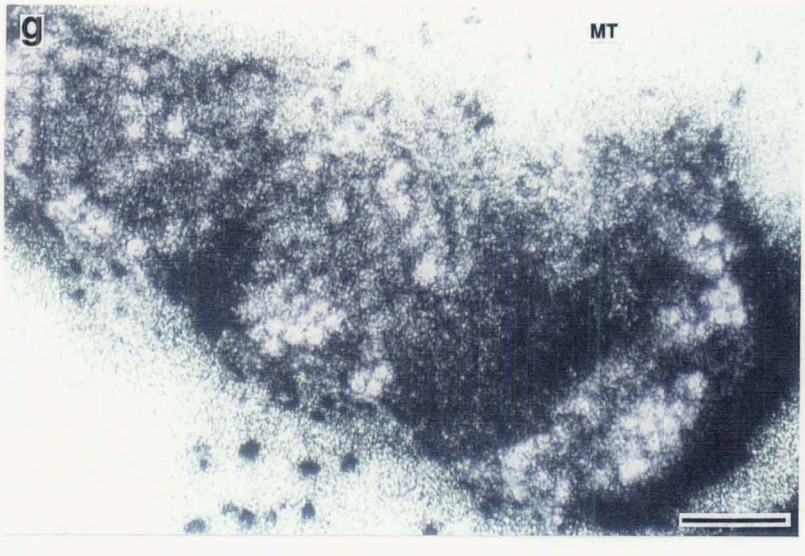
Figure 4.5. *In situ* hybridization of lobster eyestalk (lobster 1) to cDNA probe peJK1. Comparison of peJK1 probe hybridization on experimental slide (**a, b, e, f, g**), to corresponding regions on the RNase treated control sections (**c, d, h, i, j**). Probe hybridization is evident as accumulation of dense silver grains over the neurosecretory regions of the eyestalk in the whole mount (**a**). In the neurosecretory regions of the medulla externa and medulla interna (**b, f**), as well as the medulla terminalis (**g**), probe localization occurs on the nuclei of some of the cells, and only in the cytoplasm of the other neurosecretory cells. In the neurosecretory cells above the lamina ganglionaris, and retinal cells, peJK1 hybridized to the cell nuclei, and to the cytoplasm of the ommatidial cells (**e**). The corresponding region on the RNase-treated control is shown (**h**). The sections have been differentiated by staining with Haemotoxylin; cell nuclei appear as purple to pink structures, whereas cell boundaries are not as distinctive in the sections. *ME* medulla externa; *MI* medulla interna; *MT* medulla terminalis; *LG* lamina ganglionaris; *OM* ommatidia. *Bars*: a, c = 500 µm; b, d, e, f, h, and i = 100 µm; g = 50 µm.

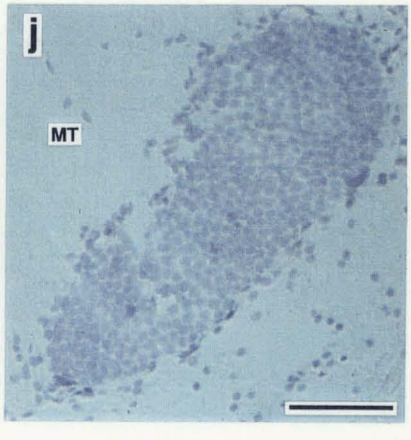

e
OM


h
OM


f
MI
MT


i
ME
MT


g
MT

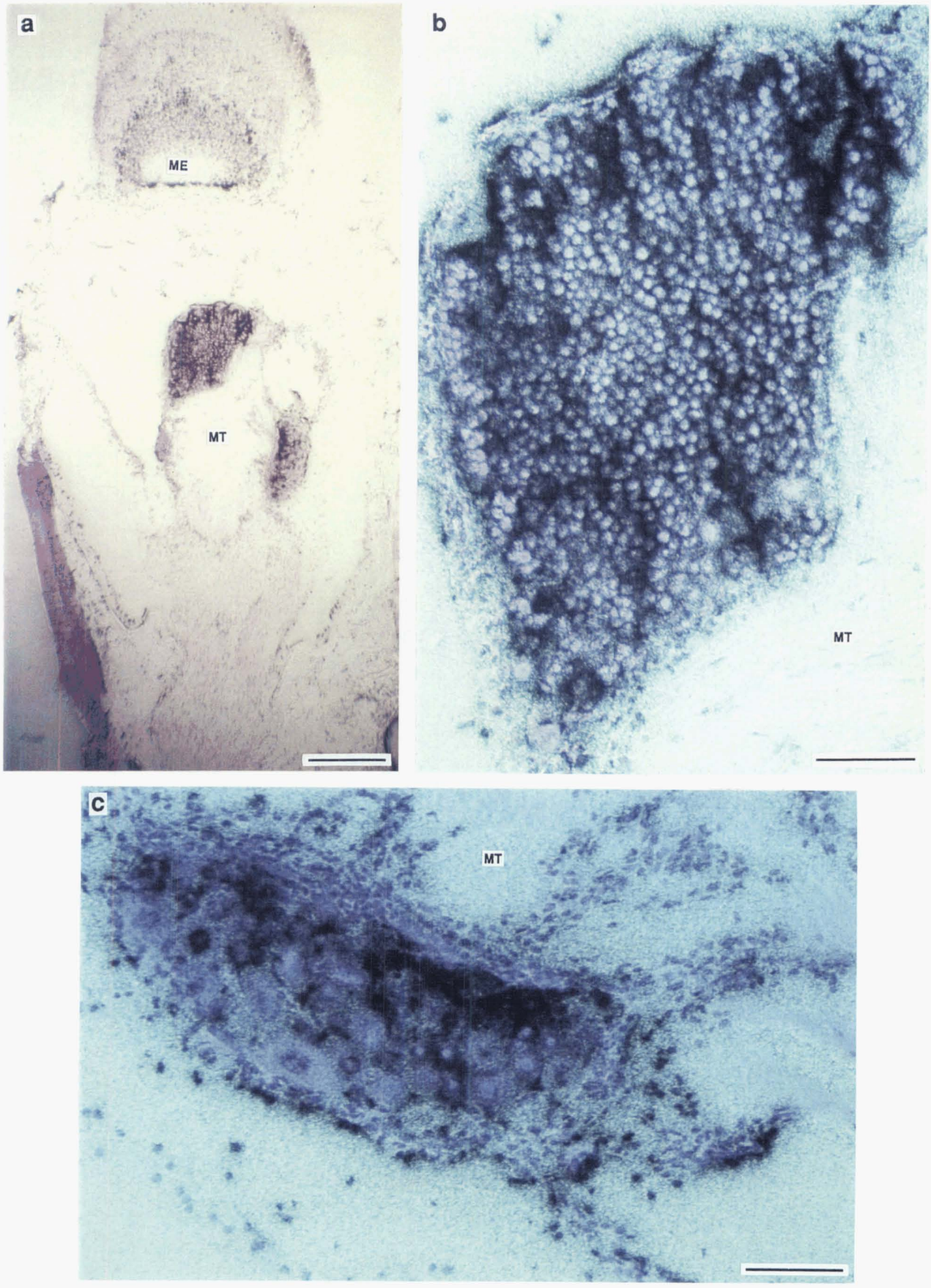

j
MT

Within the neurosecretory cells of the medulla externa, medulla interna, and medulla terminalis, three different patterns of peJK1 probe hybridization were noted; a) hybridization to the cell nucleus, b) hybridization to the cell cytoplasm, and c) hybridization to the entire cell body (perikaryon). Nuclear hybridization of the probe was seen as dense silver grains over the nucleus, whereas cytoplasmic hybridization was suggested by the ringlike pattern of silver grain accumulation around the purple stained nucleus (Fig. 4.5b, f, and g). Cell nuclei in the sections were differentiated by haematoxylin staining, and appeared pink to purple.

PeJK1 hybridization was not only confined to the neurosecretory cell groups of the medulla externa, interna and terminalis; nuclear hybridization of peJK1 was seen in the cells above the lamina ganglionaris, as well as the nuclei of the retinal cells. Cytoplasmic localization of the probe, however, was seen in the cells of the ommatidia (Fig. 4.5e). RNase digestion abolished probe binding in this region (Fig. 4.5h).

The pattern of peJK1 hybridization was also different in another section of the same eyestalk (Fig. 4.6)(ca. 400 μm from the section shown in Fig. 4.5). Based on the morphology and relative position of the structure, in comparison with published literature, a glandular structure was identified as the sinus gland (Fig. 4.6a and b). PeJK1 hybridization was confined to the cytoplasm of the cells in this structure. The accumulation of silver grains over the perikaryon in some parts of the gland is probably due to overexposure of emulsion to the probe. In the neurosecretory region of the medulla terminalis of the same section, different neurosecretory cell types were evident. The cells had large nuclei and perikarya, and the cell boundaries were distinct. PeJK1 probe hybridization in this region was patchy, with cytoplasmic hybridization occurring in some of the small and large cells, and nuclear localization in the smaller cells in this region (Fig. 4.6c). Based on the position and morphology of the cells (Smith and Naylor, 1972), these are probably neurosecretory cell type 5, found in the organ of Bellonci, in the medulla terminalis.

Figure 4.6. *In situ* hybridization of peJK1 to lobster eyestalk sections (cont.). This tissue section (**a**) was taken ca. 400 µm from the eyestalk section in Fig. 4.5. Cytoplasmic localization of peJK1 is seen in the sinus gland (**b**), and hybridization to different cell types of the medulla terminalis X-organs (**c**). Bars: a = 500 µm; b and c = 100 µm.


a
ME
MT
b
MT
c
MT

4.3.2.3 Distribution of peJK2 and peJK3 in the lobster eyestalk

Northern blot analysis showed both peJK2 and peJK3 were predominantly expressed in the lobster eyestalk. peJK2 hybridized to the neurosecretory cell groups of the medulla externa and interna, the X-organs of the medulla terminalis, and to the region above the lamina ganglionaris including the retinal and ommatidia cell nuclei (Fig. 4.7a-e). The pattern of hybridization was patchy within these areas. Nuclear hybridization of the probe was observed in some cells while other cells in the vicinity showed absence of probe hybridization (e.g. MT X-organs, Fig. 4.7d). This suggests that peJK2 is expressed only in a certain proportion of the cells in a specified time frame. Cytoplasmic localization of both peJK2 and peJK3 were also observed in the glandular structure identified as the sinus gland, as shown in the example (Fig. 4.8a and b).

Figure 4.9 shows results of peJK2 *in situ* hybridization to eyestalk sections taken from a different lobster. Both nuclear and cytoplasmic hybridization of peJK2 occurred within the neurosecretory cell groups of the medulla externa, interna, and medulla terminalis (Fig. 4.9a-d). Not all the cells show hybridization to peJK2. The interesting feature here is the presence of a distinct group of large cells near the medulla externa (Fig. 4.9b), which are likely to be cell type 6 according to the classification of Smith and Naylor (1972).

Tissue sections ca. 200-300 μm from the eyestalk sections tested with peJK1 (Fig. 4.5), were hybridized to peJK2 (Fig. 4.10). The relative sizes of the optic neuropiles are reduced as the sections were taken closer to the edge of the eyestalk; the medulla externa and interna are reduced in size but still recognizable. RNase treatment abolished peJK2 binding to mRNA transcripts within the neurosecretory regions of the medulla terminalis, and also to the nuclei of the retinal and ommatidial cells (Fig. 4.10d, e and f). Both nuclear and cytoplasmic hybridization of peJK2 occurred in the neurosecretory cells of the medulla externa, and terminalis. However, the pattern of silver grain arrangement suggests that the mRNA transcripts are closely associated with the nuclear membrane (Fig. 4.10b and c). At least two different cell types were distinguishable in the medulla terminalis X-organs (Fig. 4.10b); probe

Figure 4.7. *In situ* hybridization of peJK2 to lobster eyestalk sections, ca. 100-200 µm from the tissue section shown in Fig. 4.5. peJK2 probe localization is evident in the neurosecretory regions (stained pink) of the eyestalk section whole mount (**a**). Nuclear localization of peJK2 was shown in the neurosecretory cell groups of the medulla externa and interna (**a, c**), the medulla terminalis (**b**), the X-organs of the medulla terminalis (**d**), and the retina (**e**). Bars: a = 500 µm; b and c = 100 µm, d and e = 50 µm.

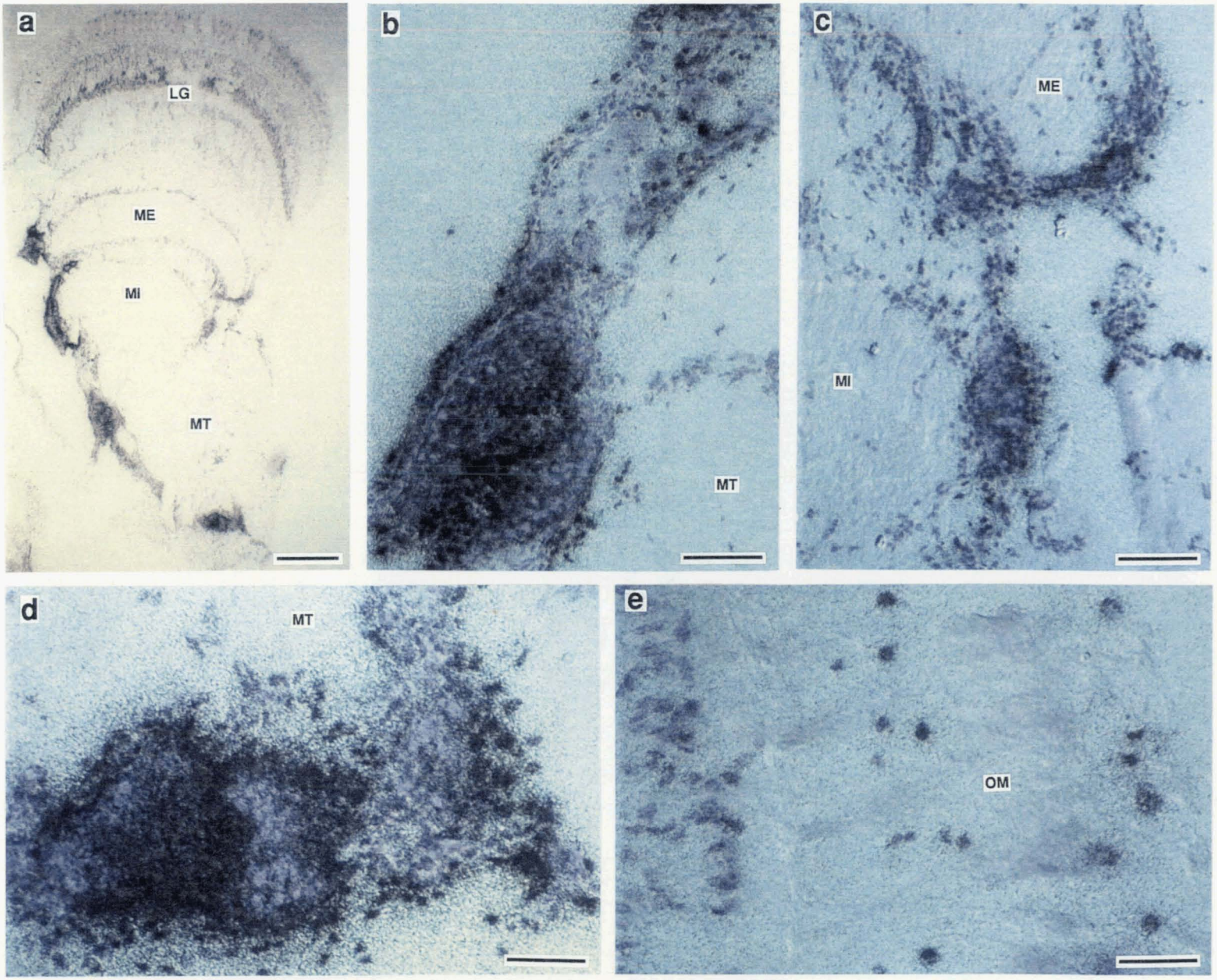

a
LG
ME
MI
MT
b
MT
c
ME
MI
d
MT
e
OM

Fig. 4.8. *In situ* hybridization of peJK3 to lobter eyestalk sections. Cytoplasmic localization of peJK3 to the sinus gland, as seen in the whole mount (**a**), and at higher magnification (**b**). This tissue section was taken from a different lobster (lobster 2). Bars: a = 500 μm; b = 100 μm.


a
LG
ME
MT

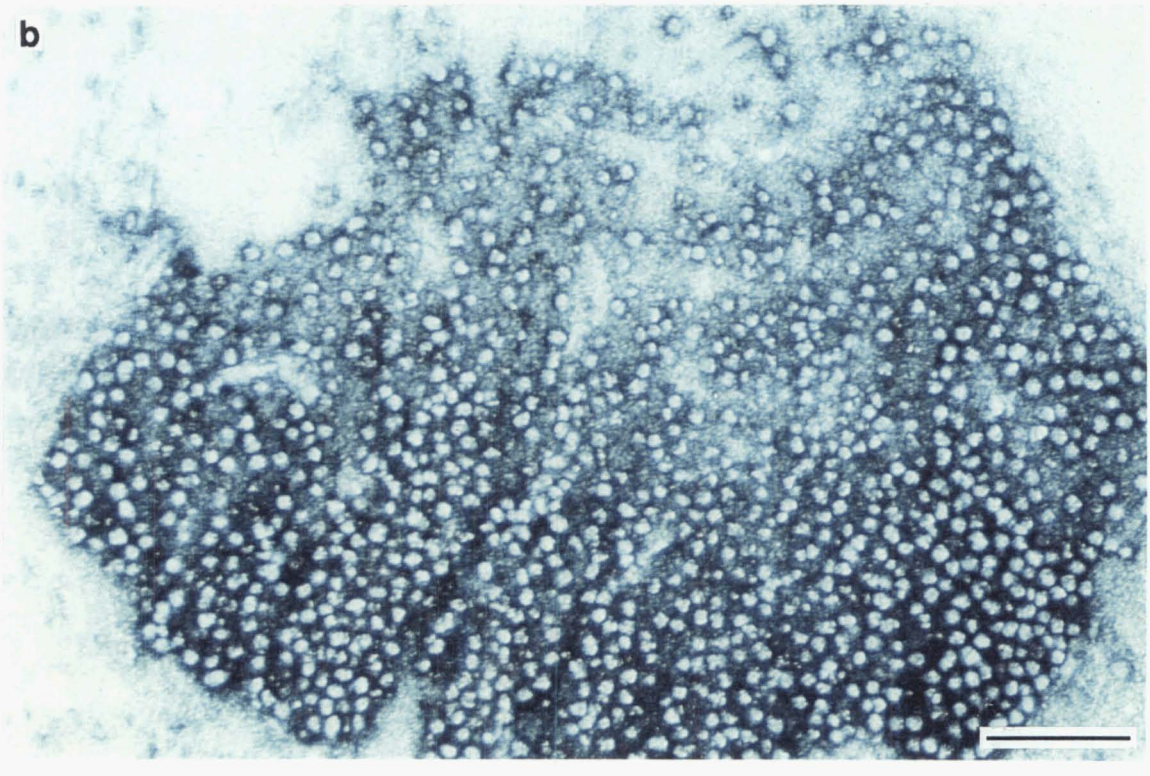

b

Figure 4.9. *In situ* hybridization of peJK2 to eyestalk sections (lobster 3). The neurosecretory regions showing probe localization are identified as the darker stained regions in the whole mount (**a**). Different cell types are evident in the neurosecretory cell groups of the medulla externa (**b**), and the medulla terminalis (**c**). Nuclear localization of peJK2 mRNAs was not found on every cell in the medulla terminalis (**c**), and the medulla terminalis X-organs (**d**). Bars: a = 500 μm; b, c, and d = 100 μm.

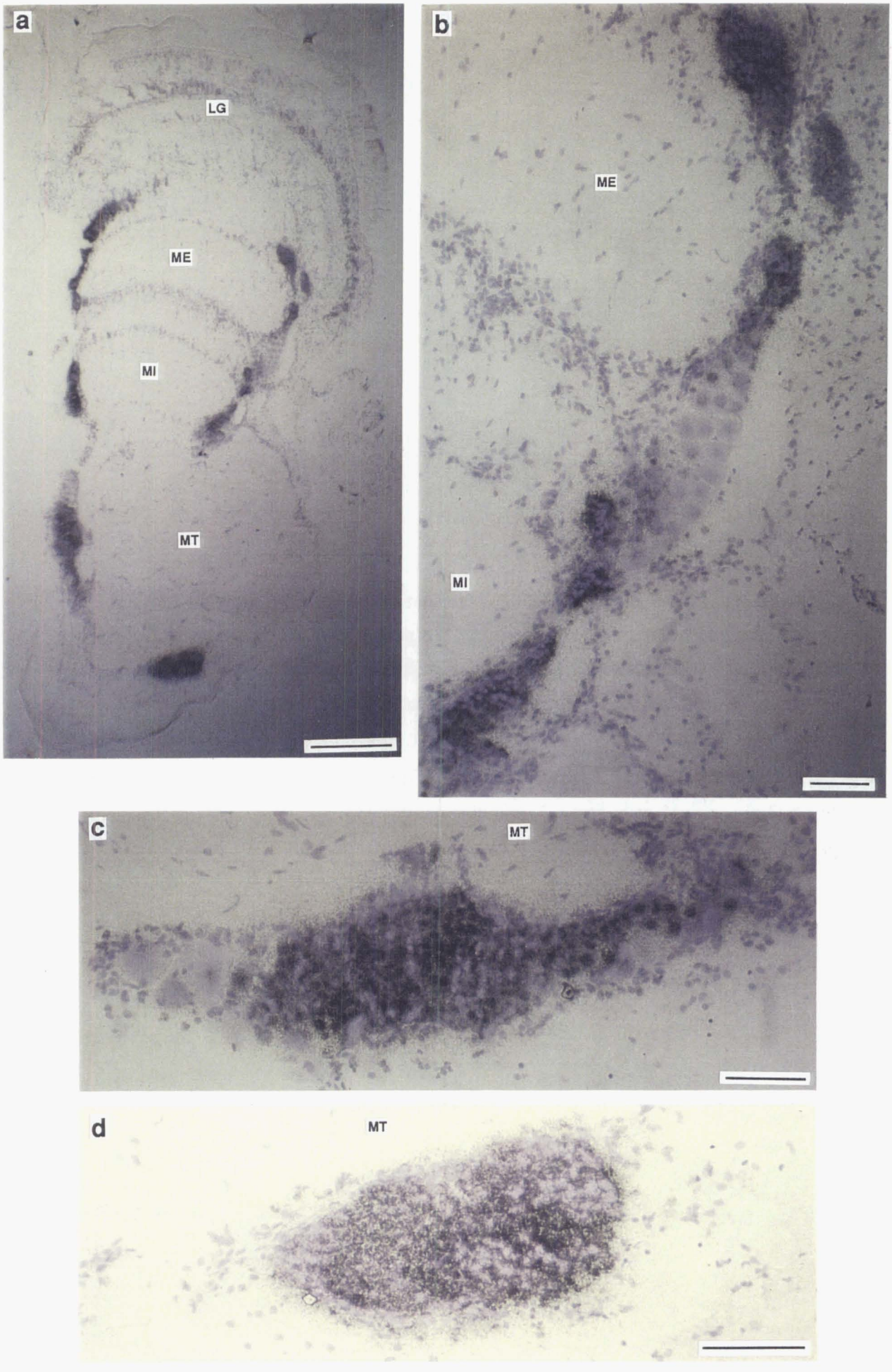
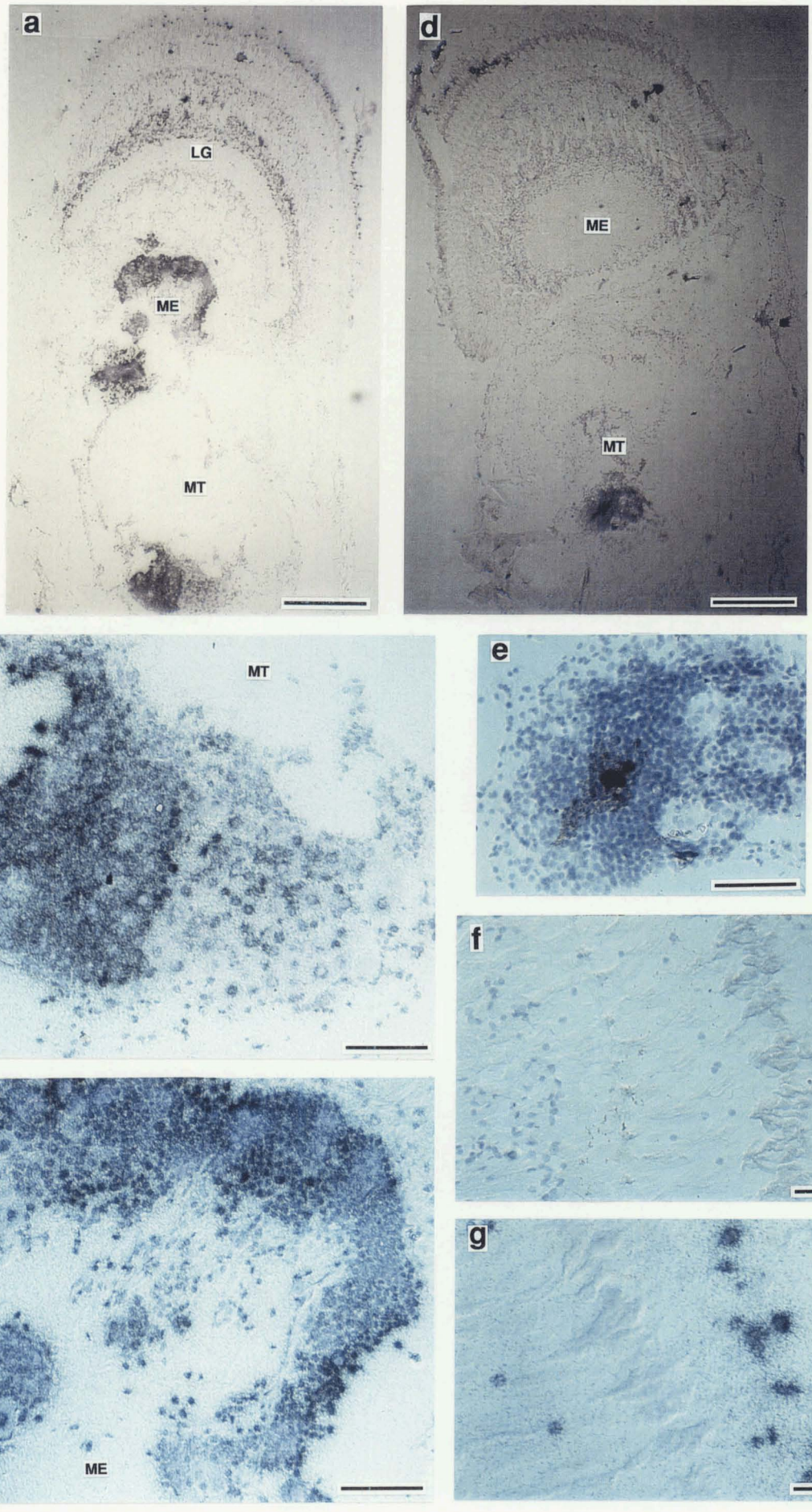

a
LG
ME
MI
MT
b
ME
MI
c
MT
d
MT

Figure 4.10. Comparison of *in situ* hybridization of peJK2 to normal and RNase treated sections (lobster 1). The nuclei of the neurosecretory cells (stained pink), are evident in the whole mount (**a**). Details of peJK2 hybridization to the medulla terminalis X-organs (**b**), neurosecretory cells of the medulla externa (**c**), and the retina (**g**). Different cell types were evident in these neurosecretory cell groups. Figures (**d, e, f**) are corresponding areas on the RNase treated control section. Bars: a and d = 500 μm; b, c, e, and f = 100 μm; g = 50 μm.


a
LG
ME
MT
d
ME
MT
b
MT
e
f
c
ME
g

hybridization occured quite uniformly over the region of the small densely packed cells, whereas in the larger cells, a gradation of silver grain density was noted, extending from the nucleus to the cytoplasm. This could suggest that the cells in this group are at different stages of the transcription cycle as both nuclear localization and cytoplasmic localization of mRNA transcripts were detected by hybridization to peJK2.

4.3.2.4 Distribution of peJK1 in the body tissues

Northern blot analysis to peJK1 demonstrated that peJK1 was also expressed in the epithelial, heart, hepatopancreas, and muscle tissue of the lobster. *In situ* hybridization of 16 µM sections of these tissues carried out simultaneously with the eyestalk section revealed the following results.

peJK1 hybridized to mRNA transcripts in the nucleus and cytoplasm of epithelial cells and myocardial (heart muscle) cells, and the nucleus of muscle cells (Fig. 4.11a-c). In both the epithelial and heart tissue, the cell boundaries were not defined by staining; the presence of silver grains all over the tissue was indicative of a low level of cytoplasmic hybridization as this was absent in the controls. In the muscle, probe localization was confined to the cell nuclei. RNase treatment of the tissue sections on adjacent slides, abolished probe hybridization, as shown by the very low density of silver grains over the entire background in the controls for each tissue type tested (Fig. 4.11e-h).

The number of nuclei which could be positively identified as showing nuclear hybridization was estimated for the three tissues. For peJK1, approximately 50% of the cell nuclei in epithelial, heart and muscle, showed nuclear hybridization to the probe. The results were based on the mean number of counts (n=20) made within a 0.04 mm^2 quadrat, for epithelial and heart tissue, and 0.155 mm^2 for muscle tissue (summarised in Table 4.4). The size of the quadrats were selected to include > 50 nuclei, and quadrats were randomly selected over one tissue section on the slide.

Figure 4.11. *In situ* hybridization of peJK1 to lobster body tissue sections. Nuclear and cytoplasmic localization of peJK1 is evident as silver grains accumulated on the nuclei and over the tissue of the l.s. epithelium (**a**), and t.s. heart (**b**). In the t.s. muscle, probe hybridization is confined to the nucleus (**c**). In the transverse sections of the gill lamellae, silver grains are distributed over the nuclei and cytoplasm of cells within the lamella (**d**). The corresponding controls for the epithelial tissue (**e**), heart (**f**), muscle (**g**) and gill (**h**), show low levels of silver grains over the tissue sections. **cy**, cytoplasm; **n**, nucleus; **m**, muscle. Bars: a, b, e and f = 20 μm; c, d, g and h = 50 μm. *l.s.* longitudinal section, *t.s* transverse section.

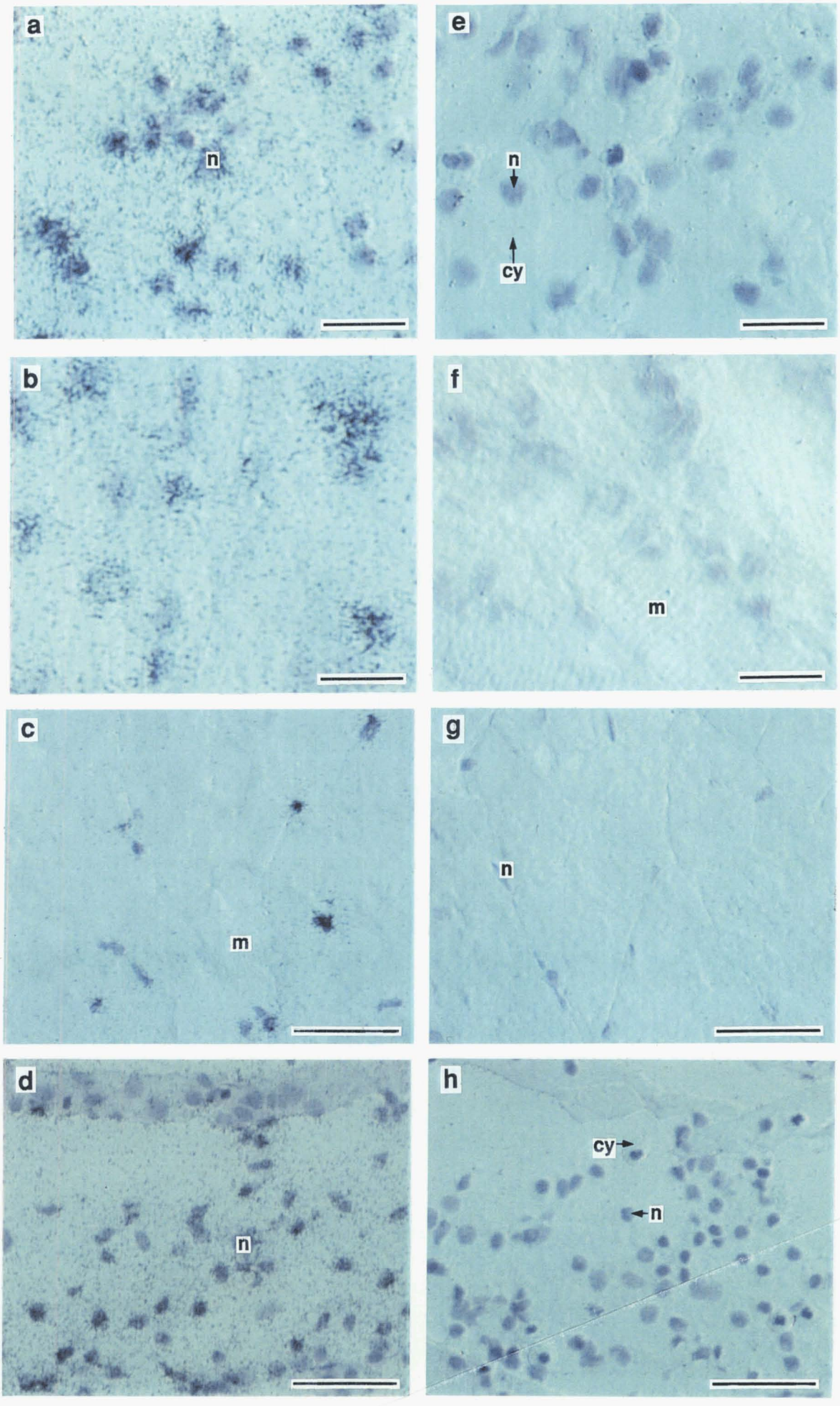

a
n
e
n
cy
b
f
m
c
m
g
n
d
n
h
cy
n

Table 4.4. Comparison of the mean number of nuclei in control and experimental sections showing hybridization to peJK1. n=20.

Tissue type	% nuclei control	% nuclei treatment	paired t-test*	quadrat area (mm^2)
heart, t.s	0	47.59 ± 8.98	$p < 0.0001$	0.04
epithelial, l.s.	0.46 ± 0.74	50.39 ± 12.35	$p < 0.0001$	0.04
muscle, t.s.	0.57 ± 1.17	54.34 ± 13.46	$p < 0.0001$	0.155

* Paired t-tests comparison of the number of nuclei showing hybridization to peJK1 in the RNase digested control and experimental tissue sections.

The level of expression of peJK1 in the heart, epithelia, and muscle were not significantly different, as shown by paired t-test comparison of the number of nuclei showing hybridization to peJK1 in the different tissues (results summarized in Table 4.5).

Table 4.5. Comparison of peJK1 nuclear hybridization in the different tissue types (from Table 4.4).

Comparison of tissue tupes	t-value	p value*
epithelia vs heart	t = 0.82	N.S.
epithelia vs muscle	t = -0.95	N.S.
heart vs muscle	t = -1.95	N.S.

* Paired t-tests comparison of the number of nuclei showing hybridization to peJK1 in the different tissue types (n=20). The t-values obtained are shown. N.S. represents non-significant p values.

Northern blot analysis failed to detect any bands within the mRNA of gill tissue. However, *in situ* hybridization revealed silver grain accumulation over the cell nuclei and cytoplasm within the gill lamellae of the experimental slide, but not in RNase treated controls (Fig. 4.11d and h).

4.3.2.5 Distribution of peJK2 in tissues other than the eyestalk

Northern blot analysis presented earlier demonstrated that peJK2 was detected predominantly in the eyestalk; relatively low amounts of peJK2 mRNA transcripts were detected in the other tissues. Both nuclear and cytoplasmic localization of peJK2 mRNA transcripts were observed for the epithelial and heart tissue, whereas in the muscle tissue, nuclear hybridization of peJK2 was observed (Fig. 4.12a-c). Approximately 53% of the myocardial cell nuclei showed probe hybridization (Table 4.6); in the epithelial tissue, however, nuclear hybridization was not as pronounced as was observed for peJK1 on the same tissue type (Table 4.4).

RNase treatment of the controls reduced the silver grain density to background levels (Fig. 4.12e-h). There were some areas where there was a high amount of silver grains on the control slides; these were confined to the blood sinus spaces of the heart, and also between the bundles of muscle fibres (in both peJK1 and peJK2). This was possibly the result of non-specific trapping of the probe in the intercellular spaces.

Table 4.6. Comparison of the average number of nuclei in control and experimental sections showing hybridization to peJK2. n=20.

tissue type	% nuclei control	% nuclei treatment	paired t-test*	quadrat area (mm^2)
heart, t.s.	0	53.53 ± 13.91	$p < 0.0001$	0.04
epithelial, l.s.	0.08 ± 0.23	28.67 ± 8.91	$p < 0.0001$	0.04
muscle, t.s.	0.30 ± 0.72	40.53 ± 10.83	$p < 0.0001$	0.155

* Paired t-tests comparison of the number of nuclei showing hybridization to peJK2 in the RNase digested control and experimental tissue sections.

When expression of peJK2 in the different tissues were compared, the number of nuclei showing nuclear hybridization was significantly different between the tissue types (Table 4.7). Nuclear hybridization of peJK2 was highest in the heart tissue, and lowest in the epithelial tissue (Table 4.6).

Figure 4.12. *In situ* hybridization of peJK2 to lobster body tissues. Nuclear and cytoplasmic hybridization is evident as silver grains accumulated over the nuclei and tissue of the epithelium (**a**), and heart (**b**). PeJK2 hybridization is confined to the nuclei of the muscle fibres (**c**), and the nuclei and cytoplasm of the cells within the gill lamella (**d**). RNase treatment of controls reduced probe binding to background levels (**e-h**). **cy**, cytoplasm; **n**, nucleus; **m**, muscle. Bars: a, b, e and f = 20 μm; c, d, g and h = 50 μm.

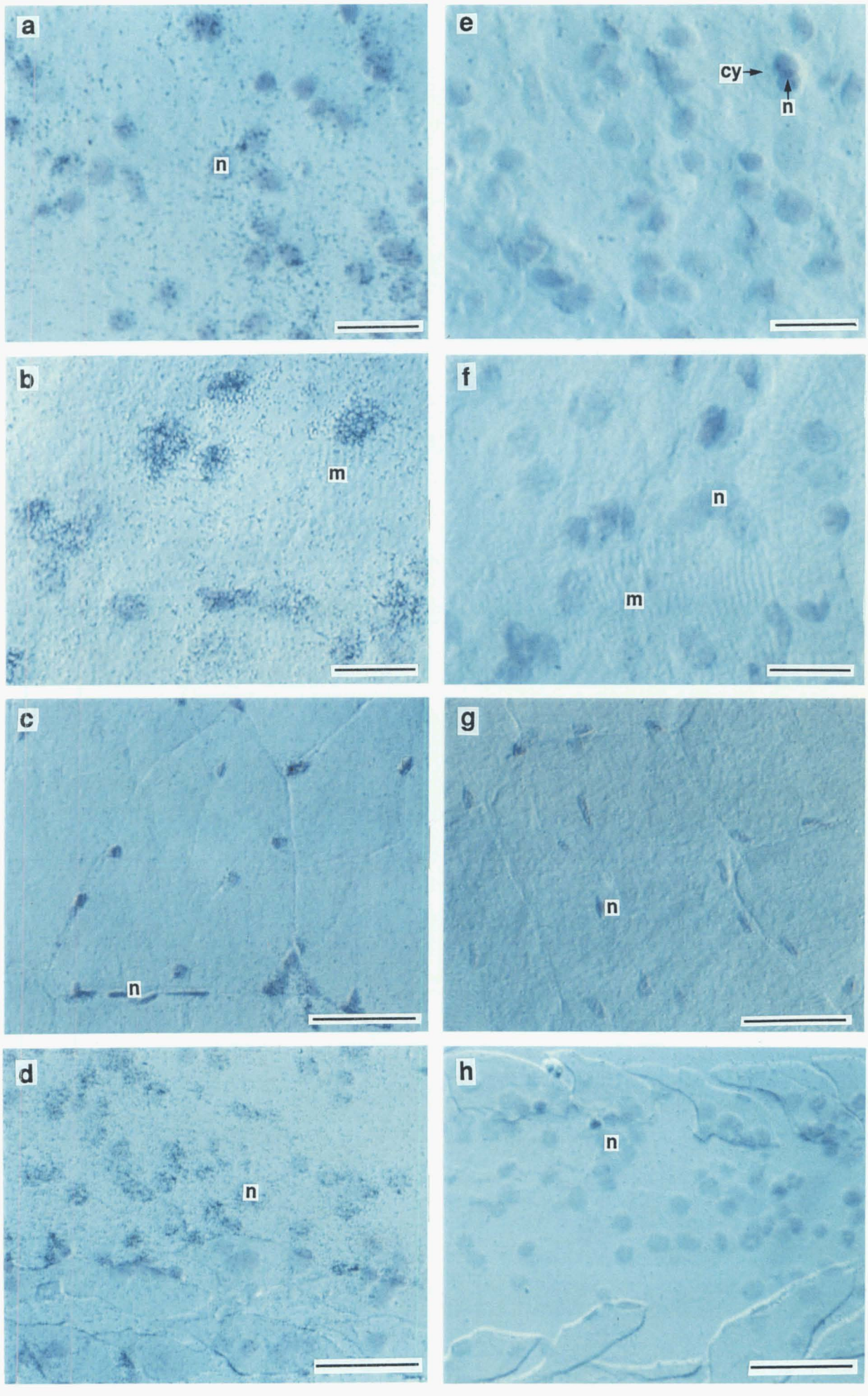

a
n
e
cy
n
b
m
f
n
m
c
n
g
n
d
n
h
n

Table 4.7. Comparison of peJK2 nuclear hybridization in the different tissue types (from Table 4.6).

Comparison of tissue tupes	t-value	p value*
epithelia vs heart	t = -7.27	p < 0.0001
epithelia vs muscle	t = -4.60	p < 0.0005
heart vs muscle	t = 3.22	p <0.005

* Paired t-tests comparison of the number of nuclei showing hybridization to peJK2 in the different tissue types (n=20). The t-values obtained are shown.

Cross-sections of gill tissue showed the same pattern of probe hybridization which was observed for peJK1, i.e., *in situ* hybridization revealed silver grain accumulation over the cell nuclei and cytoplasm within the gill lamellae of the experimental slide, but not in RNase treated controls (Fig. 4.12d and h).

The tissue sections used in hybridization to peJK1 and peJK2 were serial sections on adjacent slides. Paired t-test showed that the expression of peJK1 was significantly higher in both the epithelial and muscle tissue, compared to peJK2 (Table 4.8).

Table 4.8. Comparison of peJK1 and peJK2 nuclear hybridization in the different tissue types (from Tables 4.4 and 4.6).

Tissue type	t-value	p value*
epithelia	t = 5.45	p < 0.0001
heart	t = -1.66	N.S.
muscle	t = 3.09	p < 0.01

* Paired t-tests comparison of the number of nuclei showing hybridization to peJK1 and peJK2 in the different tissue types (n=20). The t-values obtained are shown. N.S. represents non-significant p values.

4.4 DISCUSSION

The previous chapters (Chapters II and III) described the isolation and sequence analysis of three clones, peJK1, peJK2 and peJK3, containing the putative MIH sequence, from a lobster eyestalk cDNA library. Northern blot analysis and *in situ* hybridization studies were carried out to examine the distribution and expression of these three cDNA clones, and to determine whether the expression of these sequences were similar to the published putative MIH and CHH sequences.

4.4.1 Expression of peJK1

Northern blot analysis demonstrated that peJK1 was expressed at significant levels in the epithelia, eyestalk, heart, muscle and hepatopancreas; in the first four tissues, the sizes of the major bands detected were very similar. Judging by the pattern of peJK1 hybridization to the poly(A)$^+$ RNA in these tissues, the mRNA detected either contained the peJK1 sequence or highly related sequences. Closer examination of peJK1 distribution by *in situ* hybridization revealed that peJK1 was being expressed in approximately 50% of the nuclei in the epithelia, heart and muscle tissues. Detection of peJK1 in the cytoplasm as well, suggested that mRNA transcripts were being actively transcribed, and transported into the cytoplasm.

The size of the major eyestalk mRNA was approximately 1.1 kb bigger than the actual peJK1 sequence (585 bp). The inability to obtain the full length clone could be due to low efficiency in unfolding the secondary structure of the entire mRNA prior to the reverse transcriptase reaction in cDNA synthesis. However, 10 of the 15 clones isolated from this cDNA library (10^6 pfu/µg cDNA) contained the peJK1 sequence, which would suggest that this particular sequence was abundant in the eyestalk. In addition, detection of mRNAs of varying sizes, in the other tissues suggests i) alternate splicing of a common mRNA precursor, and ii) some of these transcripts are highly related to the peJK1 sequence. The other bands were detected using high stringency conditions as well as a short exposure time (Fig. 4.1d).

The size difference between peJK1 and its corresponding mRNA could also account for the inability to detect significant sequence homology between the 947 PCR sequence, peJK2 and peJK3 sequences to the peJK1 cDNA sequence. The homologous sequence could be part of the larger mRNA precursor which was not isolated or sequenced.

4.4.2 Expression of peJK2 (and peJK3)

peJK2 detected mRNAs of a similar size in the eyestalk, epithelia and muscle extracts, suggesting that the same mRNA transcript was being produced. However, the expression of this mRNA was more prominent in the eyestalk than in the other tissues. When examined by *in situ* hybridization, the epithelial tissue showed a significantly lower proportion of the nuclei showed nuclear hybridization, compared to the heart and muscle tissues. Although there is no direct physiological evidence to suggest that peJK2 and peJK3 represent the sequence of a putative MIH neuropeptide, the predominant expression of the gene in the eyestalk and in the epithelia and muscle, while absent in the other tissues analysed, may be related to the physiological changes involved in moulting. This could suggest that peJK2 codes for a MIH-like peptide.

In crustaceans, moulting is a hormone-mediated process and the two hormones involved are MIH and ecdysterone. Previous findings suggest that in addition to suppressing ecdysteroid production by the Y-organs (Soumoff and O'Connor, 1982, reviewed in Quackenbush, 1986; Lachaise *et al.*, 1993), MIH also acts antagonistically to 20-hydroxyecdysone at the site of the epidermal tissue (Freeman and Bartell, 1976; Freeman and Costlow, 1979). It has been proposed that ecdysterone (moulting hormone) causes cellular differentiation of the epithelium by activating normally dormant genes (Stringfellow and Skinner 1988), leading to the secretion of the new cuticle. Replication of DNA in the epidermis was also found to be correlated with the moulting cycle. The peak of DNA synthesis occurred during the period of rising moulting hormone levels (Wittig and Stevenson, 1975). Moult cycle related changes in the muscle tissue, such as rates of protein synthesis, and claw muscle cyclic atrophy and restoration have also been reported (El Haj and

Houlihan, 1987; Mykles, 1992). Therefore, it is possible that MIH may be expressed in the epithelia and muscle at a lower level (basal levels) than the eyestalk, to counteract the effect of circulating ecdysteroids in the haemolymph.

4.4.3 Expression of peJK1 and peJK2 in the lobster eyestalk

Although peJK1 and peJK2 detected mRNAs in the same neurosecretory regions, the pattern of hybridization suggests that the two types of mRNAs were produced by different cells. Hybridization of adjacent slides containing serial sections of the same eyestalk, with peJK1 and peJK2 probes, produced different hybridization patterns within the medulla terminalis X-organs (Fig. 4.5g and 4.7d). Both types of results, i.e. expression of two different neuropeptides showing co-localization, and localization in different cells have been reported. *In situ* hybridization experiments by de Kleijn *et al.* (1992) demonstrated that there was a co-localization of CHH and GIH mRNA in several perikarya of the X-organs of the lobster *Homarus americanus*, but other cells reacted only with one of the two probes. However, Klein *et al.* (1993a) showed that there was no co-localization of MIH and CHH at the mRNA and protein level in the eyestalk of the crab *Carcinus maenas*.

peJK1 and peJK2 expression were also detected in the glandular structure identified as the sinus gland. However, because the identity of this structure was not verified by histological staining such as Aldehyde Fuchsin staining, the implications of these results are not discussed. The sinus gland is generally regarded as a neurohemal organ for the storage and release of hormones (Cooke and Sullivan, 1982; Fingerman, 1987).

peJK1 and peJK2 were expressed in similar regions of the lobster eyestalk. However, the distribution of peJK1 and peJK2 mRNA transcripts in the lobster eyestalk were very different from the expression of CHH and putative MIH, as detected by immunocytochemistry (Dircksen *et al.*, 1988). By use of antisera raised against purified putative MIH and CHH from the crab, *C. maenas*, Dircksen *et al.* (1988) demonstrated that CHH and MIH immunoreactivity was confined to the X-organ sinus gland system of the eyestalk in the following brachyuran species: *C.*

maenas, Liocarcinus puber, Cancer pagarus, Uca pugilator and Maja squinado. MIH-positive perikarya were localized more peripherally, and dispersed among the more numerous CHH-positive perikarya of the medulla terminalis X-organs. Both types of cells were morphologically similar and correspond to cell type 5 in *C. maenas* (Smith and Naylor, 1972). Similar results were also obtained at the mRNA level by *in situ* hybridization studies in *C. maenas* using cRNA probes (Klein *et al.*, 1993a, 1993b).

Besides differences to CHH and putative MIH localization in the eyestalk, peJK1 and peJK2 were expressed at much higher levels in the eyestalk than the putative MIH and CHH. peJK1 and peJK2 mRNAs were detected in the eyestalk after 2 h exposure, using high stringency northern conditions. However, in published results of northern blot analysis with the putative MIH and CHH cDNA sequences, MIH-like mRNAs were detected in eyestalk poly(A)$^+$ RNA (1 µg) after 8 days, using high stringency wash conditions at 68°C (Sun, 1994). Lee *et al.* (1995), using the putative MIH cDNA probe, detected a 1.4 kb band in total RNA (30 µg) from the eyestalk of the crab, *C. sapidus*. The hybridization conditions (0.1 x SSC/0.1 SDS at 55°C; exposure time was 7 days) were less stringent than those used for peJK1 and peJK2. By northern blot analysis using medium stringency conditions (0.25 x SSPE at 56°C), CHH mRNA transcripts were detected in the eyestalk after 16 h in *O. limosus*, and *H. americanus* (4 µg medulla terminalis poly(A)$^+$ RNA) (de Kleijn, *et al.*, 1994a; de Kleijn, *et al.*, 1995). The observed northern hybridization bands were much fainter than the peJK1 and peJK2 detected ones.

These differences may be due to the sensitivity of the techniques used. All of the papers examined used either immunocytochemical localization or non-radioactive *in situ* hybridization with a digoxygenin labeled probe. The immunocytochemical technique commonly used will detect proteins in the cell or gland. Because MIH and CHH are synthesized and released into the haemolymph, what is detected will be a reflection of the amount of protein present in the cell. Immunological techniques may not be sensitive enough to detect low levels of expression as the protein is not accumulated in the cell. On the other hand, *in situ* hybridization demonstrates the

presence of a specific mRNA. As mRNAs are not exported out of the cell, cells actively synthesizing the protein will therefore have a high concentration of the mRNA in the cell which allows easy detection.

The distribution of neurosecretory cells showing peJK1 and peJK2 expression was closer to the distribution of immunoreactivity to substance P (SP)-like, methionine-(Met)- and leucine-(Leu)-enkephalin-like, and FMRFamide-like peptides in the eyestalk of the fiddler crab *Uca pugilator* (Fingerman *et al.*, 1985). However, peJK1 and peJK2 cDNA sequences showed no significant homology to the rat enkephalin sequence (Yoshikawa *et al.*, 1984), and the mollusc FMRF amide sequences (*L. stagnalis*, Linacre *et al.*, 1990; *Aplysia*, Schaefer *et al.*, 1985). Both [Leu]- and [Met]enkephalin have been purified from thoracic ganglia of the shore crab *C. maenas*; automated gas-phase sequencing revealed a primary structure identical with that of enkephalins from vertebrates (Lüschen *et al.*, 1991). A similar pattern of distribution was also demonstrated immunocytochemically for the red pigment concentrating hormone (RPCH) in the eyestalk of the crayfish, *Procambarus clarkii* (Rodríguez-Sosa, *et al.*, 1994). However, there was no significant alignment between the deduced amino acids of the RPCH precursor (Klein *et al.*, 1992) and that of peJK1 and peJK2.

Based on the expression of peJK1, peJK2 and peJK3, as characterized by northern blot analysis and *in situ* hybridization studies, peJK1 is not likely to contain the putative MIH sequence as it was expressed in all the tissues tested. However, because peJK1 is produced both in the neurosecretory cells of the eyestalk, as well as other tissues, peJK1 may be coding for a regulatory peptide, in the sense of a local cell-to-cell communicator, based on the relative abundance of mRNAs detected in the other tissues. peJK2 and peJK3 could contain a MIH-like related sequence as it is expressed predominantly in the eyestalk.

CHAPTER V

General Discussion and Conclusions

Moulting in crustaceans is a complex process which affects every aspect of the animal's physiology. One of the factors which has been implicated in the control of moulting is the eyestalk neuropeptide, MIH.

5.1 Overview of the experimental approach used

The aim of this thesis was to isolate and characterise the putative MIH sequence from the lobster *J. edwardsii* using an approach commonly used for searching for homologous sequences in unrelated species. Peptides of the vasopressin-oxytocin family were shown to be related to the MIH in the crab *Cancer antennarius*, furthermore, vasopressin- and neurophysin-like peptides have been detected in the eyestalks of the prawn *Palaemon serratus* (Van Herp and Bellon-Humbert, 1982), and have been found in several invertebrates (Cruz *et al.*, 1987; Proux *et al.*, 1987; van Kesteren *et al.*, 1992; Reich, 1992; Oumi *et al.*, 1994). Thus the use of DNA probes derived from the conserved sequences of the cloned rat vasopressin gene may allow the isolation of the MIH-related neuropeptide genes from the lobster eyestalk.

Although Southern hybridization demonstrated the presence of vasopressin-like gene sequences in the lobster genome, the preliminary search for the vasopressin-like gene sequence by genomic library screening was unsuccessful. This could have been due to the large proportion of intron sequence (75%) in the rat vasopressin gene (Schmale *et al.*, 1983), or low homology to the lobster gene. Subsequently PCR was used to amplify the vasopressin-like gene sequence. Vasopressin is only a nine amino acid residue peptide, however, the gene sequence encoding the peptide is more complex (described in Chapter I). PCR primer sequences were based on the

conserved regions of the rat vasopressin gene encoding part of the vasopressin hormone (primer 1) and the conserved region of the neurophysin (primer 2).

Using these primers, a 947 PCR product was isolated. Sequence analysis revealed a putative protein coding region with intron sequence. To verify that this PCR product was really coding for mRNA, northern blot analysis was carried out. Northern blot analysis using the 947 PCR product as a probe showed that the putative coding region was represented in the poly(A)$^+$ RNA species isolated from the various lobster tissues but was expressed predominantly in the eyestalk. *In situ* hybridization studies showed the expression of the gene in the neurosecretory regions of the eyestalk. There was also circumstantial evidence to suggest that the expression was related to the moulting cycle. The relative amount of mRNA detected by the 947 PCR fragment was significantly reduced in the eyestalk sections taken from a lobster in the premoult stage as compared to the intermoult stage. Taken together, these results suggested that the 947 PCR product was related to the vasopressin-like peptide and might have a physiological role in relation to moulting.

The 947 PCR product was subsequently used as a probe to screen the cDNA and genomic libraries. Two phage clones were retrieved from the lobster genomic DNA library, and these were subcloned as plasmids p2A and p66 (7.4 kb and 8.0 kb inserts respectively). Sequencing of these plasmids however, is still incomplete (Chapter III). One of these clones, plasmid p66, was used to screen the cDNA libraries constructed from mRNA isolated from lobster eyestalks (Chapter III).

The cDNA libraries were initially screened with the 947 PCR product but I failed to isolated any positive clones. This is contrary to expectation as *in situ* hybridization and northern analysis clearly demonstrated that the 947 PCR product is related to some mRNAs in the lobster eyestalk. Thus, the failure to retrieve any positive clones using the 947 PCR product as a probe could be due to the relative proportions of coding to non-coding sequences within the 947 PCR sequence.

The genomic clone, plasmid p66, was then used for the primary screening of the amplified cDNA libraries, and the 947 PCR sequence was used for secondary screening. Using this approach, three cDNA clones, peJK1, peJK2 and peJK3 were isolated.

The questions arising are:

(i) Are any of these three genes related to MIH?

(ii) Are they related to vasopressin?

5.2 Sequence comparison

The amino acid sequences of putative MIH isolated from the lobster, *H. americanus*, and the shore crab, *C. maenas*, have been reported (Chang *et al.*, 1990; Webster, 1991). Based on these sequences, PCR was used to generate probes which were used in cDNA library screening for clones encoding the putative MIH neuropeptide. To date, cDNA sequences encoding the putative MIH have been reported for *C. maenas*, *P. vannamei*, and *C. sapidus* (Klein *et al.*, 1993b; Sun, 1994; and Lee *et al.*, 1995). However, the identity of the definitive MIH has not been established as the precise significance and function of these neuropeptides as a moult-inhibitor *in vivo* has not yet been elucidated.

Comparison of peJK1, peJK2 and peJK3 cDNA sequences with published putative MIH cDNA sequences from other species (Klein *et al.*, 1993b; Sun, 1994; Lee *et al.*, 1995) showed that peJK1, peJK2 and peJK3 were quite different. peJK1 shared 44-48% sequence homology, and peJK2 and peJK3 shared 46-51% sequence homology with the published putative MIH cDNA sequences. Within the homologous regions, there were no significant contiguous nucleotides which were identical.

Furthermore, based on amino acid sequence alignments and comparisons, the predicted protein structure of peJK1, peJK2 and peJK3, were quite different from the structure of published putative MIH. The six invariant cysteine residues commonly

found in the putative MIH (Chang *et al.*, 1990; Klein *et al.*, 1993b; Sun, 1994; Lee *et al.*, 1995) were absent from the deduced amino acid sequences of peJK1, peJK2 and peJK3. In the related CHH eyestalk neuropeptide, the six cysteine residues have been shown to contribute to the overall structure of the peptide by the formation of disulphide bridges between the cysteine residues (Kegel *et al.*, 1989; Yasuda *et al.*, 1994; Huberman *et al.*, 1995).The absence of these cysteine residues in the predicted peptides of peJK1, peJK2 and peJK3, suggests that these cDNA clones do not code for peptides similar to published putative MIH peptides.

The MIH in the crab *Cancer antennarius* was shown to be related both structurally and functionally to the vasopressins, lysine vasopressin in particular (Mattson and Spaziani, 1985a). Sinus gland extracts of *Cancer magister* also show a peak on HPLC which has both MIH-like activity, and a retention time similar to that of authentic Lys-vasopressin (Mattson, unpublished; as reported in Webster and Keller, 1989). These results strongly suggest that *Cancer* MIH is similar if not identical to Lys–vasopressin.

However, Lys- and Arg-vasopressin-like MIH have not been found in *Carcinus maenas* using HPLC (unpublished results; reported in Webster and Keller, 1989), but immunopositive structures were detected in the eyestalk of *Cancer pagurus* when *C. maenas* MIH antisera were used to detect MIH (Dircksen *et al.*, 1988). Thus, a substance identical or very similar to the MIH of *C. maenas* appears to be present in *Cancer pagurus.*

A search on the GenBank and EMBL datalibraries showed that peJK1 had similar sequences to the *O. limosus* CHH-A cDNA sequence (51.5% sequence identity over 400 bp overlap). However, because peJK1 is expressed in significant levels in the epithelia, eyestalk, heart, and muscle; it is unlikely to be related to CHH. In other species, the CHH neuropeptide has been isolated from the eyestalk (Kegel *et al.*, 1989; Kegel *et al.*, 1991; Huberman *et al.*, 1993) and only recently in the ventral nervous system, as detected by northern blot analysis (de Kleijn *et al.*, 1995).

5.3 Tissue distribution of MIH

The two tissues, eyestalk and epithelia, are implicated in relation to MIH and moulting (Soumoff and O'Connor, 1982; Freeman and Bartell, 1976; Freeman and Costlow 1979; Webster, 1993; reviewed in Lachaise *et al.*, 1993). The distribution of peJK1 and peJK2 were shown to be different by northern blot analysis. peJK1 was expressed in the epithelia, eyestalk, heart, and muscle, whereas peJK2 and peJK3 were expressed predominantly in the eyestalk, and at lower levels in the epithelia and muscle tissues. peJK2 and peJK3 are possibly allelic forms of the same gene, based on the 96.6% sequence identity shared between the two sequences.

In situ hybridization studies showed that both peJK1 and peJK2 were expressed in similar neurosecretory regions of the eyestalk but different cells could be producing peJK1 and peJK2 mRNA transcripts. These results differ from previous reports where MIH production/secretion has been found in a subset of cells of the medulla terminalis, and sinus gland, using both immunocytochemical techniques (Dircksen *et al.*, 1988), and non-radioactive *in situ* hybridization studies (Klein *et al.*, 1993a, 1993b). However, Webster (1986) demonstrated that in addition to the medulla terminalis, both the medulla externa and medulla interna had a small effect *in vitro* on the inhibition of ecdysteroid synthesis by *Carcinus* Y-organs, *in vitro*.

Webster (1993) reported the presence of putative MIH receptors on the Y-organs. However, the results of small amounts of MIH receptors found on the epidermis was not published (as reported in Lachaise *et al.*, 1993). Recently, Sun (1995) detected a MIH-like gene transcript in the brain of the shrimp, *P. vannamei*, using the reverse transcriptase-polymerase chain reaction (RT-PCR), a more superior method for detecting small amounts of mRNA. Subsequent cloning and sequencing of the PCR products showed that the MIH-like gene transcript of the brain shares 98% sequence identity with that of the eyestalk (Sun, 1994). *In situ* hybridization experiments showed that the MIH-like mRNA was localized exclusively in the X-organ complex of the medulla terminalis in the eyestalk, whereas in the brain, the MIH-like gene transcript was detected in regions including the neurosecretory cells, giant cells and lateral cell bodies. Based on the data, Sun (1995) suggested that the

MIH-like peptide could have a specific function in nervous system activity other than its classic moult-inhibiting function.

The wide distribution pattern of peJK2 as detected by northern blot analysis and *in situ* hybridization, as compared to the narrow distribution of other major eyestalk neuropeptides (CHH/MIH/VIH) in other species (Dircksen *et al.*, 1988; Kallen and Meusy, 1989; Tensen *et al.*, 1991a; de Kleijn *et al.*, 1992, Klein *et al.*, 1993a, 1993b; Laverdure *et al.*, 1994) may be a reflection of the techniques used. For example, fixation of the tissues for frozen sections, the use of a much longer probe, and the use of high energy [α-^{32}P] dCTP labelled probe as compared to a non-radioactive digoxigenin cRNA probe; all of which may have increased the sensitivity of the *in situ* hybridization technique.

Traditionally, immunocytochemical techniques have been used to detect peptides within the eyestalk (Van Herp and Bellon-Humbert, 1982; Van Deijnen *et al.*, 1985, Dircksen *et al.*, 1988; Kallen and Meusy, 1989). This method detects the specific protein present in the cell. Based on pulse chase experiments on the isolated X-organ-sinus gland of *Cardisoma carnifex*, Stuenkel *et al.* (1991) demonstrated that peptides synthesized in the X-organs were transported to the sinus gland (neurohemal organ) within a day; the rate of axonal transport of the neurosecretory granules were between 0.1-0.2 mm/h. Thus, immunocytochemical methods may detect only low levels of MIH in the eyestalk at any one time during the premoult stage. The *in situ* hybridization technique detects mRNAs which are not exported out of the cell, and therefore could be a more sensitive method for the detection of low levels of gene expression in tissues such as the epidermis and muscle.

Based on my observation on the distribution of the cDNA sequences in the gill tissues, northern blot analysis was less sensitive than the *in situ* hybridization technique in demonstrating whether or not a particular gene was being expressed. By northern blot analysis, no mRNA was detected in the gill tissue extract. However, *in situ* hybridization demonstrated the presence of such sequences in the nuclei of the gill tissue.

The previous studies on putative MIH have examined for tissue specific expression using northern blot analysis. Hybridization to the eyestalk mRNA and total RNA was detected after a 7 to 8 day long exposure period (Sun, 1994; Lee *et al.*, 1995). This result might lead one to the conclusion that MIH is not present in any other tissue. Therefore, the presence of putative MIH in other tissues may not have been examined by a more sensitive method.

The distribution of peJK1 detected mRNA transcripts in the eyestalk, however, is similar to the distribution of the met- and leu-enkephalin-like peptides detected in the eyestalk of the fiddler crab *Uca pugilator* (Fingerman *et al.*, 1985). Homology searches have not revealed any significant similarities to mammalian enkephalins; the peJK1 sequence is incomplete as it is derived from a much larger transcript.

In mammals, the enkephalins which are pentapeptides, are a group of opioid peptides, whose physiological actions can be described as those of neuromodulators-neurotransmitters, hormones, or both (reviewed in Gubler, 1987). In crustaceans, a neuromodulatoy role of [Leu]enkephalin in the regulation of blood glucose levels has been reported; the release of the CHH was shown to be inhibited by synthetic Leu-enkephalin (Rothe *et al.*, 1991). The involvement of [Met]enkephalin in pigment regulation (i.e. adaption to black background) in the fiddler crab (Kulkarni and Fingerman, 1987) as well as the participation of the opioids in the escape responses in the land crab *Gecarcinus lateralis* (Martinez *et al.*, 1988) indicates a crustacean opioid system (as cited in Lüschen *et al.*, 1991). Both [Leu]- and [Met]enkephalin have been purified from thoracic ganglia of the shore crab *C. maenas* (Lüschen *et al.*, 1991). It is possible that peJK1 may be related to this peptide, or have a similar role as a neuromodulator, or local cell to cell communicator. The vertebrate hormone vasopressin also has this function (Richter, 1987).

5.4 Temporal expression of MIH

Although there is no direct evidence to suggest that peJK2 and peJK3 code for a putative MIH, there is indirect evidence linking the expression of peJK2 and peJK3 to the moult cycle. In section 2.3.6 (Chapter II), the expression of the 947 PCR

sequence in the eyestalk was shown to be significantly reduced in the premoult lobster eyestalk as compared to the intermoult lobster. The differences observed between the premoult and the intermoult eyestalk could probably be attributed to the expression of peJK2 and peJK3. Using northern blot analysis, the major eyestalk mRNA species detected by peJK2 and peJK3, was also the major mRNA species in the eyestalk detected by the 947 PCR product. Taken together, these results suggest that peJK2 and peJK3 could code for a MIH-like peptide.

CONCLUSIONS

Using the approach detailed earlier, several cDNA and genomic sequences have been retrieved from the lobster *J. edwardsii*. Both peJK2 and peJK3 appeared to be allelic forms of a single gene which is expressed predominantly in the eyestalk. peJK1, however, was expressed significantly in the eyestalk as well as other tissues. At this stage it is not possible to unequivocally state that any of these sequences are the putative MIH gene. peJK2 and peJK3, however, code for a major eyestalk neuropeptide with a signal peptide. Sequence analysis of these clones showed that they shared 43 to 47% sequence identity to the rat vasopressin gene, and about 50% sequence identity to the MIH/CHH sequences of other crustaceans.

The biological functions or significance of the proteins encoded by peJK1, peJK2 and peJK3 can be investigated by subcloning the full length cDNA sequences of peJK1, peJK2 and peJK3 into a plasmid expression vector in the correct orientation. The proteins expressed in a coupled transcription translation cell free system can be isolated and the physical properties characterized. The possible roles of these putative proteins in the moult cycle can then be examined *in vitro* by incubating the proteins with cultured Y-organs and measuring the effect on ecdysteroid production by the Y-organs. Inhibition of ecdysteroid production would support a MIH-like role for the proteins. The effects of injections of purified proteins on circulating ecdysteroid levels in the whole animal can also be studied.

The expression of the peJK genes in relation to the moult cycle can also be examined by RT-PCR detection of the peJK gene transcripts and/or by *in situ* hybridization of eyestalk sections taken from lobsters at determined stages of the moult cycle.

The relationship between the cDNA clones peJKs and the 947 PCR product has yet to be clarified. Sequence comparisons so far failed to identify the homologous regions. Complete sequencing of the genomic clones may provide insights into the relationship between the cDNA clones and the 947 PCR product, and will provide insights into the identity of the various mRNA products, and evidence for alternative splicing mechanisms in the lobster.

ACKNOWLEDGEMENTS

I thank Dr. Frank Sin who has been an excellent supervisor and teacher and under whom I have learnt a lot. I also thank Dr. Liz Allison for her input and time into my work. Melanie North and Jan McKenzie's help was invaluable in getting this thesis together. Terry Williams also did a painstaking job on the photography, and overall, the technical support was excellent. I would like to express my gratitude to all the people in this department who have made my time here a pleasure, and have encouraged me along the way. I appreciate the support and friendship that I have received since I have been here. I was supported financially by the Roper Science scholarship from the University of Canterbury.

I am very thankful and grateful to my family and friends who have supported me throughout this period, Daddy and Mummy for giving me leave to pursue whatever plan I had, Kilistina for your encouragement all the way through, and the saints at Happy Church, who have supported and encouraged me all these years. Finally, thank you God for your provision, your strength, faithfulness and the dream that you placed in my heart.

REFERENCES

Aiken, D.E. (1980) Molting and growth. *In* The biology and management of lobsters: Physiology and Behaviour 1. J.S. Cobb and B.F. Phillips, eds. Academic Press, New York. pp. 91-163.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology.* **215**, 403-410.

Amara, S.G. (1985) Peptide diversity from exon choice: RNA processing regulation in the neuroendocrine system (Review). *Molecular and cellular endocrinology* **42**, 191-205.

Anderson, M.L.M. and Young, B.D. (1985) Quantitive filter hybridization. *In* Nucleic acid hybridisation. A practical approach. B.D. Hames and S.J. Higgins, eds. IRL Press, Oxford. pp. 73-111.

Archer, R. (1981) Molecular evolution of the polypeptide hormones: the neurohypophysial hormone-neurophysin model. Neuropeptides Biochemical and Physiological Studies . R.P. Miller, eds. Churchhill Livingstone, Great Britain. pp. 240-262.

Archer, R. and Chauvet, J. (1988) Structure, processing and evolution of the neurohypophysial hormone-neurophysin precursors. *Biochimie* **70**, 1197-1207.

Benton, W.D. and Davis, R.W. (1977) Screening lambda gt recombinant clones by hybridization to single plaques *in situ. Science* **196**, 180-182.

Birnstiel, M.L., Busslinger, M. and Strub, K. (1985) Transcription termination and 3' proceesing: The end is in site! *Cell* **41**, 349-359.

Bliss, D.E. and Welsh, J.H. (1952) The neurosecretory system of brachyuran crustacea. *Biological Bulletin marine biological laboratory Woods Hole* **103**, 157-169.

Bloch, B., Popovici, T., Le Guellec, D., Normand, E., Chouham, S., Guitteny, A.F., Bohlen, P. (1986) *In situ* hybridization histochemistry for the analysis of gene expression in the endocrine and central nervous system tissues: a 3-year experience. *Journal of Neuroscience Research* **16**, 183-200.

Bonwick, G.A., Fielden, P.R. and Davies, D.H. (1991) Hepatic metallothionein levels in roach (*Rutilus ruticlus*L.) continously exposed to water-borne cadmium. *Comparative Biochemistry Physiology* **99C**, 119-125.

Bullock, T.H. and Horridge, G.A. (1965) Structure and function in the nervous system of invertebrates. 1 and 2 W.H. Freeman and Company, Netherlands.

Carlisle, D.B., and Passano, L.M. (1953) The X-organ of Crustacea. *Nature, London* **171**, 1070-1071.

Chang, E.S. (1985) Hormonal control of molting in decapod crustacea. *American Zoologist* **25**, 179-185.

Chang, E.S. (1995) Physiological and biochemical changes during the molt cycle in decapod crustaceans: an overview. *Journal of Experimental Marine Biology and Ecology* **193**, 1-14.

Chang, E.S., Bruce, M.J. and Newcomb, R.W. (1987) Purification and amino acid composition of a peptide with molt-inhibiting activity from the lobster, *Homarus americanus*. *General and Comparative Endocrinology* **65**, 56-64.

Chang, E.S., Bruce, M.J. and Tamone, S.L. (1993) Regulation of crustacean molting: A multi-hormonal system. *American Zoologist* **33**, 324-329.

Chang, E.S., Prestwich, G.D. and Bruce, M.J. (1990) Amino acid sequence of a peptide with both molt-inhibiting and hyperglycemic activities in the lobster *Homarus americanus*. *Biochemical and Biophysical Research Communications*. **171**(2), 818-826.

Chesselet, M.-F. (1990) In situ hybridization histochemistry. CRC Press Inc., Florida, U.S.A.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* **162**, 156-159.

Clark, J.M. (1988) Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Research* **16**(20), 9677-9687.

Cooke,I.M., and Sullivan, R.E. (1982). Hormones and neurosecretion. *In* The biology of Crustacea, Vol. 3, Neurobiology: structure and function. D.E. Bliss, H.L. Atwood and D.C. Sandeman eds. Academic Press, New York. pp. 205-290.

Cruz, L.J., de Santos, V., Zafaralla, G.C., Ramilo, C.A., Zeikus, R., Gray, W.R. and Olivera, B.M. (1987) Invertebrate vasopressin/oxytocin homologs. *The Journal of Biological Chemistry*. **262**(33), 15821-15824.

de Kleijn, D.P.V., Coenen, A.J.M., Laverdue, A.M., Tensen, C.P. and Van Herp, F. (1992) Localization of mRNAs encoding crustacean hyperglycemic hormone and gonad inhibiting hormone in the x-organ sinus gland complex of the lobster *Homarus americanus*. *Neuroscience* **51**, 121-128.

de Kleijn, D.P.V., de Leeuw, E.P.H., van den Berg, M.C., Martens, G.J.M. and van Herp, F. (1995) Cloning and expression of two mRNAs encoding structurally different crustacean hyperglycemic hormone precursors in the lobster *Homarus americanus*. *Biochimica et Biophysica Acta* **1260**, 62-66.

de Kleijn, D.P.V., Janssen, K.P.C., Martens, G.J.M. and Van Herp, F. (1994a) Cloning and expression of two crustacean hyperglycemic-hormone mRNAs in the eyestalk of the crayfish *Orconectes limosus*. *European Journal of Biochemistry*. **224**, 623-629.

de Kleijn, D.P.V., Sleutels, F.J.G.T., Martens, G.J.M. and Van Herp, F. (1994b) Cloning and expression of mRNA encoding prepro-gonad-inhibitng hormone (GIH) in the lobster *Homarus americanus*. *FEBS Letters* **353**, 255-258.

de Kleijn, D.P.V. and Van Herp, F. (1995) Molecular biology of neurohormone precursors in the eyestalk of crustacea (review). *Comparative Biochemistry Physiology* **112B**(4), 573-579.

De Loof, A. and Schoofs, L. (1990) Homologies between the amino acid sequences of some vertebrate peptide hormones and peptides isolated from invertebrate sources. *Comparative Biochemistry Physiology* **95B**(3), 459-468.

Dircksen, H., Webster, S.G. and Keller, R. (1988) Immunocytochemical demonstration of the neurosecretory systems containing putative moult-inhibiting hormone and hyperglycemic hormone in the eyestalk of brachyuran crustaceans. *Cell and Tissue Research* **251**, 3-12.

Djangmah, J.S. and Grove, D.J. (1970) Blood and hepatopancreas copper in *Crangon vulgaris* (Fabricius). *Comparative Biochemistry Physiology* **32**, 733-745.

Durand, J.B. (1956) Neurosecretory cell types and their secretory activity in the crayfish. *Biological Bulletin* **111**, 62-76.

El Haj, A.J. and Houlihan, D.F. (1987) *In vitro* and *in vivo* protein synthesis rates in a crustacean muscle during the moult cycle. *Journal of Experimental Biology* **127**, 413-426.

Engel, D.W. (1987) Metal regulation and molting in the blue crab, *Callinectes sapidus:* copper, zinc and metallothionein. *Biological Bulletin* **172**, 69-82.

Fickett, J.W. (1982) Recognition of protein coding regions in DNA sequences. *Nucleic Acids Research* **10**(17), 5303-5318.

Fingerman, M. (1987) The endocrine mechanisms of crustaceans. *Journal of Crustacean Biology* **7**(1), 1-24.

Fingerman, M., Hanumante, M.M., Kulkarni, G.K., Ikeda, R. and Vacca, L.L. (1985) Localization of substance P-like, leucine-enkephalin-like, methionine-enkephalin-like, and FMRFamide-like immunoreactivity in the eyestalk of the fiddler crab *Uca pugilator. Cell and Tissue Research* **241**, 473-477.

Finney, M. (1994) Molecular cloning of PCR products. *In* Current protocols in molecular biology, vol 2. F.M. Ausubel, R. Brent, R.E Kingston, D.D. Moore, J.G Seidman, J.A. Smith, and K. Struhl, eds. John Wiley and Sons Inc., USA.

Freeman, J.A. and Bartell, C.K. (1976) Some effects of molt-inhibiting hormone and 20-hydroxyecdysone upon molting in the grass shrimp, *Palaemonetes pugio*. *General and Comparative Endocrinology* **28,** 131-142.

Freeman, J.A. and Costlow, J.D. (1979) Hormonal control of apolysis in barnacle mantle tissue epidermis, *in vitro*. *The Journal of Experimental Zoology* **210,** 333-345.

Gabe, M. (1966) Neurosecretion. Pergamon Press, Oxford.

Gubler, U. (1987) Enkephalin genes. *In* Molecular cloning of hormone genes. J.F. Habener ed. The Humana Press Inc., Clifton, New Jersey, pp. 229-276.

Habener, J.F. (1987) Gene structure and regulation. *In* Molecular cloning of hormone genes. J.F. Habener, ed. The Humana Press Inc., Clifton, New Jersey. pp. 11-51.

Haurowitz, F. (1963) The chemistry and function of proteins. Academic Press, New York.

Holmes, D.S. and Quigley, M. (1981) A rapid boiling method for the preparation of bacterial plasmids. *Analytical Biochemistry* **114,** 193-197.

Huberman, A. and Aguilar, M.B. (1989) A neuropeptide with molt-inhibiting hormone activity from the sinus gland of the mexican crayfish *Procambarus bouvieri* (Ortmann). *Comparative Biochemistry Physiology* **93B**(2), 299-305.

Huberman, A., Aguilar, M.B., Brew, K., Shabanowitz, J. and Hunt, D.F. (1993) Primary structure of the major isomorph of the crustacean hyperglycemic hormone (CHH-I) from the sinus gland of the mexican crayfish *Procambarus bouvieri* (Ortmann): Interspecies comparison. *Peptides* **14,** 7-16.

Huberman, A., Aguilar, M.B. and Quackenbush, L.S. (1995) A neuropeptide family from the sinus gland of the Mexican crayfish, *Procambarus bouvieri* (Ortmann). *Aquaculture* **135**, 149-160.

Huynh, T.V., Young, R.A. and Davis, R.W. (1985) Constructing and screening cDNA libraries in lambda gt10 and lambda gt 11. *In* DNA cloning. A practical approach 1. D.M. Glover, eds. IRL Press, Oxford. pp. 49-78.

Ivell, R. and Richter, D. (1984) Structure and comparison of the oxytocin and vasopressin genes from rat. *Proceedings of the National Academy of Science USA*. **81**, 2006-2010.

Jendrisak, J., Young, R.A. and Engel, J.D. (1987) Cloning cDNA into lambda gt10 and Lambda gt11. *In* Guide to molecular cloning techniques 152. S.L. Berger and A.R. Kimmel, eds. Academic Press, U.S.A. pp. 359-371.

Kaiser, K. and Murray, N.E. (1985) The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries. *In* DNA cloning. A practical approach 1. D.M. Glover, eds. IRL Press, Oxford. pp. 1-47.

Kallen, J. and Meusy, J.-J. (1989) Do the neurohormones VIH (vitellogenesis inhibiting hormone) and CHH (crustacean hyperglycemic hormone) of crustaceans have a common precursor? Immunolocalization of VIH and CHH in the X-organ sinus gland complex of the lobster, *Homarus americanus*. *Invertebrate Reproduction and Development* **16**, 43-52.

Kegel, G., Reichwein, B., Tensen, C.P. and Keller, R. (1991) Amino acid sequence of crustacean hyperglycemic hormone (CHH) from the crayfish, *Orconectes limosus:* Emergence of a novel neuropeptide family. *Peptides* **12**, 909-913.

Kegel, G., Reichwein, B., Weese, S., Gaus, G., Peter-Katalinic, J. and Keller, R. (1989) Amino acid sequence of the crustacean hyperglycemic hormone (CHH) from the shore crab *Carcinus maenas*. *Federation of European Biochemical Societies Letters* **255**, 10-14.

Keller, R. (1992) Crustacean neuropeptides: Structures, functions and comparative aspects (Review). *Experientia* **48**, 439-448.

Klein, J.M., de Kleijn, D.P.V., Keller, R. and Weidemann, W.M. (1992) Molecular cloning of crustacean pigment dispersing hormone precursor. *Biochemical and Biophysical Research Communications*. **189**(3), 1509-1514.

Klein, J.M., de Kleijn, D.P.V., Hunemeyer, G., Keller, R. and Weidemann, W. (1993a) Demonstration of the cellular expression of genes encoding molt-inhibiting hormone and crustacean hyperglycemic hormone in the eyestalk of the shore crab *Carcinus maenas*. *Cell and Tissue Research* **274**, 515-519.

Klein, J.M., Mangerich, S., de Kleijn, D.P.V., Keller, R. and Weidemann, W.M. (1993b) Molecular cloning of crustacean putative molt-inhibiting hormone (MIH) precursor. *Federation of European Biochemical Societies Letters* **334**(1), 139-142.

Knowles, F.G., and Carlisle, D.B. (1956) Endocrine control in the Crustacea. *Biological Reviews* **31**, 396-473.

Koenen, M. (1989) Recovery of DNA from agarose gels using liquid nitrogen. *Trends in Genetics* **5**,137.

Kulkarni, G.K., and Fingerman, M. (1987) Distal retinol pigment of the fiddler crab, *Uca pugilator:* release of the dark-adapting hormone by methionine enkephalin and FMRF-amide. *Pigment Cell Research* **1**, 51-56.

Lachaise, F., Le Roux, A., Hubert, M. and Lafont, R. (1993) The molting gland of crustaceans: Localization, activity, and endocrine control (A review). *Journal of Crustacean Biology* **13**(2), 198-234.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

Land, H., Grez, M., Ruppert, S., Schmale, H., Rehbein, M., Richter, D. and Schôtz, G. (1983) Deduced amino acid sequence from the bovine oxytocin-neurophysin I precursor cDNA. *Nature* **302**, 342-344.

Land, H., Schôtz, G., Schmale, H. and Richter, D. (1982) Nucleotide sequence of cloned cDNA encoding bovine arginine vasopressin-neurophysin II precursor. *Nature* **295**, 299-303.

Laverdure, A.-M., Carette-Desmoucelles, C., Breuzet, M. and Descamps, M. (1994) Neuropeptides and related nucleic acid sequences detected in peneid shrimps by immunochemistry and molecular hybridizations. *Neuroscience* **60**(2), 569-579.

Lee, F.-J.S., Moss, J. and Lin, L.-W. (1992) A Simplified procedure for hybridization of RNA blots. *Biotechniques* **13**(6), 130-132.

Lee, K.J., Elton, T.S., Bej, A.K., Watts, S.A. and Watson, R.D. (1995) Molecular cloning of a cDNA encoding putative molt-inhibitibg hormone from the blue crab, *Callinectes sapidus. Biochemical and Biophysical Research Communications*. **209**(3), 1126-1131.

Lehninger, A.L., Nelson, D.L. and Cox, M.M. (1993) Principles of Biochemistry. Worth Publishers, New York.

Leitch, A.R., Schwarzacher, T., Jackson, D. and Leitch, I.J. (1994) *In situ* hybridization: a practical guide. BIOS Scientific Publishers Limited, Oxford, U.K.

Lewin, B. (1994) Genes V. Oxford University and Cell Press, USA.

Linacre, A., Kellett, E., Saunders, S., Bright, K., Benjamin, P.R. and Burke, J.F. (1990) Cardioactive neuropeptide Phe-Met-Arg-Phe-NH2 (FMRFamide) and novel related peptides are encoded in multiple copies by a single gene in the snail *Lymnaea stagnalis. Journal of Neuroscience* **10**, 412-419.

Liu, Y.-C. and Chou, Y.-C. (1990) Formaldehyde in formaldehyde /agarose gel may be eliminated without affecting the electrophoretic separation of RNA molecules. *Biotechniques* **9**(5), 558-560.

López, A.J. (1995) Developmental role of transcription factor isoforms generated by alternative splicing. *Developmental Biology* **172**, 396-411.

Lüschen, W., Buck, F., Willig, A. and Jaros, P.P. (1991) Isolation, sequence analysis, and physiological properties of enkephalins in teh nervous tissue of the shore crab *Carcinus maenas* L. *Proceedings of the National Academy of Science USA.* **88**, 8671-8675.

Mancillas, J.R., McGinty, J.F., Selverston, A.I., Karten, H. and Bloom, F.E. (1981) Immunocytochemical localization of enkephalin and substance P in retina and eyestalk neurones of lobster. *Nature* **293**, 576-578.

Maniatis, T. (1991) Mechanisms of alternative pre-mRNA splicing. *Science* **251**, 33-34.

Maniatis, T., Fritsch, E.F., Sambrook, J. (1982) Molecular cloning: a laboratory manual. Cold Spring Harbour Laboratory, New York.

Marchuk, D., Drumm, M., Saulino, A. and Collins, F.S. (1991) Construction of T-vectors, a rapid and general system for the direct cloning of unmodified PCR products. *Nucleic Acids Research* **19**(5), 1154.

Martinez, E.A., Murray, M., Leung, M.K., and Stefano, G.B. (1988) Evidence for dopaminergic and opioid involvement in the regulation of locomotor activity in the land crab *Gecarcinus lateralis*. *Comparative Biochemistry Physiology* **90C**, 89-93.

Mattson, M.P. and Spaziani, E. (1985a) Functional relations of crab molt-inhibiting hormone and neurohypophysial peptides. *Peptides* **6**, 635-640.

Mattson, M.P. and Spaziani, E. (1985b) Characterization of molt-inhibiting hormone (MIH) action on crustacean Y-organ segments and dispersed cells in culture and a bioassay for MIH activity. *The Journal of Experimental Zoology* **236**, 93-101.

Mattson, M.P. and Spaziani, E. (1985c) Cyclic AMP mediates the negative regulation of Y-organ ecdysteroid production. *Molecular and cellular endocrinology* **42**, 185-189.

McGinnis, W., Shermoen, W. and Beckendorf, S.K. (1983) A transposible element inserted just 5' to a drosophila glue protein gene alters gene expression and chromatin structure. *Cell* **34**, 75-84.

McKeown, M. (1992) Alternate mRNA splicing. *Annual Review of Cell Biology* **8**, 133-155.

Midsukami, M. (1979) Physiological salines: the formulae for animals from protozoa to vertebrata. Keigaku, Tokyo.

Mount, S.M. (1982) A catalogue of splice junction sequences. *Nucleic Acids Research* **10**(2), 459-472.

Mykles, D.L. (1992) Getting out of a tight squeeze-enzymatic regulation of claw muscle atrophy in molting. *American Zoologist* **32**, 485-494.

Naya, Y., Kishida, K., Sugiyama, M., Murata, M., Miki, W., Ohnishi, M. and Nakanishi, K. (1988) Endogenous inhibitor of ecdysone synthesis in crabs. *Experientia* **44**, 50-52.

Naya, Y., Ohnishi, M., Ikeda, M., Miki, W. and Nakanishi, K. (1989) What is molt-inhibiting hormone? The role of an ecdysteroidogenesis inhibitor in the crustacean molting cycle. *Proceedings of the National Academy of Science USA.* **86**, 6826-6829.

Oumi, T., Ukena, K., Matsushima, O., Ikeda, T., Fujita, T., Minakata, H. and Nomoto, K. (1994) Annetocin: An oxytocin-related peptide isolated from the earthworm, *Eisenia foetida. Biochemical and Biophysical Research Communications.* **198**(1), 393-399.

Pardue, M.L. (1985) *In situ* hybridisation. *In* Nucleic acid hybridisation. A practical approach. B.D. Hames and S.J. Higgins, eds. IRL Press, Oxford. pp. 179-202.

Passano, L.M. (1953) Neurosecretory control of molting in crabs by X-organ-sinus gland complex. *Physiologia Comparata et Oecologia* **3**, 155-189.

Proux, J.P., Miller, C.A., Li, J.P., Carney, R.L., Girardie, A., Delaage, M. and Schooley, D.A. (1987) Identification of an arginine vasopressin-like diuretic hormone from *Locusta migratoria. Biochemical and Biophysical Research Communications.* **149**(1), 180-186.

Quackenbush, L.S. (1986) Crustacean endocrinology, a review. *Canadian Journal of Fisheries and Aquatic Science* **43**, 2271-2281.

Quackenbush, L.S. and Herrnkind, W.F. (1983) Partial characterization of eyestalk hormones controlling moult and gonadal development in the spiny lobster *Panulirus argus*. *Journal of Crustacean Biology* 3(1), 34-44.

Radtke, F., Georgiev, O., Müller, H.-P., Brugnera, E. and Schaffner, W. (1995) Functional domains of the heavy metal-responsive transcription regulator MTF-1. *Nucleic Acids Research* 23(12), 2277-2286.

Rainbow, P.S. (1988) The significance of trace metal concentrations in decapods. *In* Aspects of Decapod Crustacean Biology. Zoological Society of London symposia 59 eds. A.A. Fincham, P.S. Rainbow. Clarendon Press, Oxford. pp 291-313.

Rangarao, K. (1965) Isolation and partial characterization of moult-inhibiting hormone of the crustacean eyestalk. *Experientia* **21**, 593-594.

Reich, G. (1992) A new peptide of the oxytocin/vasopressin family isolated from the nerves of the cephalopod *Octopus vulgaris*. *Neuroscience Letters* **134**, 191-194.

Richter, D. (1983) Vasopressin and oxytocin are expressed as polyproteins. *Trends In Biochemical Sciences* **8**, 278-281.

Richter, D. (1987) Oxytocin and vasopressin genes. Expression and structure. *In* Molecular cloning of hormone genes. J.F. Habener, ed. The Humana Press Inc., Clifton, New Jersey. pp. 173-206.

Rodríguez-Sosa, L., Calderón, J., Becerra, E. and Aréchiga, H. (1994) Regional distribution and immunocytological localization of red pigment concentrating hormone in the crayfish eyestalk. *General and Comparative Endocrinology* **95**, 443-456.

Rosen, K.M. and Villa-Komaroff, L. (1990) An alternative method for the visualization of RNA in formaldehyde agarose gels. *Focus* 12(2), 23-24.

Rosenberg, S.M. (1987) Improved *in vitro* packaging of lambda DNA. *Methods in Enzymology* **153**, 95-103.

Rothe, H., Lüschen, W., Asken, A., Willig, A., and Jaros, P.P. (1991) Purified crustacean enkephalin inhibits release of hyperglycemic hormone in the crab *Carcinus maenas* L. *Comparative Biochemistry Physiology* **99C**, 57-62.

Rudolph, P.H. and Spaziani, E. (1990) Distribution of serotonergic neurons in the eyestalk and brain of the crab *Cancer antennarius*. *Comparative Biochemistry Physiology* **97C**, 241-245.

Ruppert, S., Scherer, G. and Schütz, G. (1984) Recent gene conversion involving bovine vasopressin and oxytocin precursor genes suggested by nucleotide sequence. *Nature* **308**, 554-557.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular cloning. A laboratory manual. 1,2,3 Cold Spring Harbor Laboratory Press, USA.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Science USA*. **74**(12), 5463-5467.

Santos, M.A. (1991) An inproved method for the small scale preparation of bacteriophage DNA based on phage precipitation by zinc chloride. *Nucleic Acids Research* **19**(19), 5442.

Schaefer, M., Picciotto, M.R., Kreiner, T., Kaldany, R.-R., Taussig, R. and Scheller, R.H. (1985) Aplysia neurons express a gene encoding multiple FMRFamide neuropeptides. *Cell* **41**, 457-467.

Schmale, H., Heinsohn, S. and Richter, D. (1983) Structural organization of the rat gene for the arginine vasopressin-neurophysin precursor. *EMBO Journal* **2**(5), 763-767.

Silhavy, T.J., Berman, M.L. and Enquist, L.W. (1984) Experiments with gene fusion. Cold Spring Harbour Laboratory, New York.

Skinner, D.M. (1985) Interacting factors in the control of the crustacean molt cycle. *American Zoologist* **25**, 275-284.

Smith, G. and Naylor, E. (1972) The neurosecretory system of the eyestalk of *Carcinus maenas* (Crustacea: Decapoda). *Journal of Zoology London* **166,** 313-321.

Soumoff, C. and O'Connor, J.D. (1982) Repression of Y-organ secretory activity by molt inhibiting hormone in the crab *Pachygrapsus crassipes. General and Comparative Endocrinology* **48,** 432-439.

Soyez, D. and Kleinholz, L.H. (1977) Molt-inhibiting factor from the crustacean eyestalk. *General and Comparative Endocrinology* **31,** 233-242.

Stringfellow, L.A. and Skinner, D.M. (1988) Molt-cycle correlated patterns of synthesis of integumentary proteins in the land crab *Gecarcinus lateralis. Developmental Biology* **128,** 97-110.

Stuenkel, E., Gillary, E. and Cooke, I. (1991) Autoradiographic evidence that transport of newly synthesized neuropeptides is directed to release sites in the X-organ-sinus gland of *Cardisoma carnifex. Cell and Tissue Research* **264,** 253-262.

Sun, P.S. (1994) Molecular cloning and sequence analysis of a cDNA encoding a molt-inhibiting hormone-like neuropeptide from the white shrimp *Penaeus vannamei. Molecular Marine Biology and Biotechnology* **3**(1), 1-6.

Sun, P.S. (1995) Expression of the molt-inhibiting hormone-like gene in the eyestalk and the brain of the white shrimp *Penaeus vannamei. Molecular Marine Biology and Biotechnology* **4**(3), 262-268.

Taylor, W.R. (1987) Protein structure prediction. *In* Nucleic acid and protein sequence analysis. A practical approach. M.J. Bishop and C.J. Rawlings, eds. IRL Press, Oxford. pp. 285-322.

Tensen, C.P., Coenen, T. and Van Herp, F. (1991a) Detection of the mRNA encodong crustacean hyperglycemic hormone (CHH) in the eyestalk of the crayfish *Orconectes limosus* using non-radioactive *in situ* hybridization. *Neuroscience Letters* **124,** 178-182.

Tensen, C.P., De Kleijn, D.P.V. and Van Herp, F. (1991b) Cloning and sequence analysis of cDNA encoding two crustacean hyperglaycemic hormones from the lobster *Homarus americanus. European Journal of Biochemistry.* **200**, 103-106.

Towner, P. (1993) Purification of DNA. *In* Essential Molecular Biology. A practical approach 1. D. Rickwood and B.D. Hames, eds. IRL Press, New York. pp. 47-68.

van Deijnen, J.E., Vek, F. and Van Herp, F. (1985) An immunocytochemical study of the optic ganglia of the crayfish *Astacus leptodactylus* (Nordmann 1842) with antisera against biologically active peptides of vertebrates and invertebrates. *Cell and Tissue Research* **240**, 175-183.

von Heijne, G. (1986) A new method for predicting signal sequencee cleavage sites. *Nucleic Acids Research* **14**(11), 4683-4690.

von Heijne, G. (1987) Sequence analysis in molecular biology. Treasure Trove or Trivial Pursuit. Academic Press Inc., USA.

van Herp, F. and Bellon-Humbert, C. (1982) Localisation immunocytochimique de substances apparentées à la neurophysine et à la vasopressine dans le pédoncule oculaire de Palaemon serratus Pennant (Crustacé Décapode Natantia). *Comptes Rendus des Seances de L' Acadamie des Sciences Ser 3 Sciences de la vie* **295C**, 97-102.

van Kesteren, R.E., Smit, A.B., Dirks, R.W., de With, N.C., Geraerts, W.P.M. and Joosse, J. (1992) Evolution of the vasopressin/oxytocin superfamily: Characterization of a cDNA encoding a vasopressin-related precursor, preproconopressin, from the mollusc *Lymnaea stagnalis. Proceedings of the National Academy of Science USA.* **89**, 4593-4597.

Webster, S.G. (1986) Neurohormone control of ecdysteroid biosynthesis by *Carcinus maenas* Y-organs *in vitro*, and preliminary characterization of the putative molt-inhibiting hormone (MIH). *General and Comparative Endocrinology* **61**, 237-247.

Webster, S.G. (1991) Amino acid sequence of putative moult-inhibiting hormone from the crab *Carcinus maenas. Proceedings of the Royal Society London B* **244**,

247-252.

Webster, S.G. (1993) High-affinity binding of putative moult-inhibiting hormone (MIH) and crustacean hyperglycaemic hormone (CHH) to membrane-bound receptors on the Y-organ of the shore crab *Carcinus maenas*. *Proceedings of the Royal Society London B* **251**, 53-59.

Webster, S.G. and Keller, R. (1986) Purification, characterisation and amino acid composition of the putative moult-inhibiting hormone (MIH) of *Carcinus maenas* (Crustacea, Decapoda). *Journal of Comparative Physiology B* **156**, 617-624.

Webster, S.G. and Keller, R. (1988) Physiology and biochemistry of crustacean neurohormonal peptides. Neurohormones in invertebrates . M.C. Thorndyke and G.J. Goldsworthy, eds. Cambridge University Press, Great Britain. pp. 173-196.

Webster, S.G. and Keller, R. (1989) Molt inhibiting hormone. Ecdysone, from chemistry to mode of action . J. Koolman, eds. Thieme Medical Publishers, Inc., New York. pp. 211-216.

Weidemann, W., Gromoll, J. and Keller, R. (1989) Cloning and sequence analysis of cDNA for precursor of a crustacean hyperglycemic hormone. *Federation of European Biochemical Societies Letters* **257**(1), 31-34.

Williams, J.G. and Mason, P.J. (1985) Hybridization in the analysis of RNA. *In* Nucleic acid hybridization. B.D. Hames and S.J. Higgins, eds. IRL Press, Oxford. pp. 139-160.

Wicks, R.J. (1986) RNA molecular weight determination by agarose gel electrophoresis using formaldehyde as denaturant: comparison of RNA and DNA molecular weight markers. *International Journal of Biochemistry* **18**, 277.

Wittig, K. and Stevenson, J.R. (1975) DNA synthesis in the crayfish epidermis and its modification by ecdysterone. *Journal of Comparative Physiology* **99**, 279-286.

Yoshikawa, K., Williams, C. and Sabol, S.L. (1984) Rat brain preproenkephalin mRNA: cA cloning, primary structure, and distribution in the central nervous system. *Journal of Biological Chemistry* **259**, 14301-14308.

Yasuda, A., Yasuda, Y., Fujita, T. and Naya, Y. (1994) Characterization of crustacean hyperglycemic hormone from the crayfish (*Procambarus clarkii*): Multiplicity of molecular forms by stereoinversion and diverse functions. *General and Comparative Endocrinology* **95**, 387-398.

Zeleny, C. (1905) Compensatory regulation. *The Journal of Experimental Zoology* **2**, 1-102.

Zeller, R. (1989) *In situ* hybridization and immunohistochemistry. *In* Current Protocols in Molecular Biology 2. F.M. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl, eds. John Wiley and Sons Inc., U.S.A.

Zuckerkandl, E. (1960) Hémocyanine et cuivre chez un crustacé décopode, dans leur rapports avec le cycle d'intermue. *Annls Inst. océanogr., Monaco* **38**, 1-122.

Appendix 1

A1.1 Bacterial and phage strains

(1) LE 392 = *E. coli* F⁻ *hsd*R514 (r_{K-}, m_{K+}) *sup*E44 *sup*F58 *lac*Y1 or ∆(*lac*IZY)6 *gal*K2 *gal*T22 *met*B1 *trp*R55

(2) λgt10 = λ*srI*λ1° *b*527 *srI*λ3° *imm*^{434}(*srI*434^{+}) *srI*λ4° *srI*λ5°

(3) SMR 10 = *E. coli* λ*cos*2 ∆*B* *xis*1 *red*3 *gam*am210 c*Its*857 *nin*5 *S*am7/λ

(4) KW 251 = F⁻, *sup*E44, *gal*K2, *gal*T22, *met*B1, *hsd*R2, *mrc*B1, *mcr*A, [*arg*A81:Tn10}, *rec*D1014

A1.2 Media

Plates: Add 15g bacto-agar per liter of media

top agar: 7g of bacto-agar/ agarose per liter of media

LB medium (Luria-Bertani Medium) (per liter)

Bacto-tryptone	10g
Bacto-yeast extract	5g
NaCl	10g

TB medium (tryptone broth, λ medium) (per liter)

Bacto-tryptone	10g
NaCl	5g

After autoclaving, add 10 ml of 1 M $MgSO_4$

NZCYM medium (per liter)

NZ amine (Casein hydrolysate enzymatic)	10g
NaCl	5g
bacto-yeast extract	5g
casamino acids	1g
$MgSO_4.7H_20$	2g

Phage storage media (SM) (per liter)

0.1 M NaCl, 8.11 mM $MgSO_4.7H_20$, 50 mM Tris-HCl (pH 7.5), 0.01% gelatin

A1.3 Preparation of plating cells (*E. coli* cells used for the growth of bacteriophage).

Plating cells are prepared by innoculating a single colony of bacteria in 50 ml of TB media supplemented with 10 mM $MgSO_4$, 0.2% maltose; the culture is grown overnight at 30°C. The minimal media, and temperature used keeps the cells as close to log phase as possible in the morning. Lambda requires Mg++ for stability and adsorption to *E.coli*, whereas maltose induces the *lamB* product, thereby increasing the number of λ receptors on the cell surface (Silhavy, *et al.*, 1984). Cells are kept at 4°C, and used in the same day.

APPENDIX 2

A2.1 Isolation of the putative MIH gene sequence from the lobster genomic DNA library

The construction and screening of the lobster genomic libraries were carried out according using standard procedures (Maniatis *et al.*, 1982; Kaiser and Murray, 1985; Sambrook *et al.*, 1989). High molecular weight lobster DNA was partially digested with *Sau*3A, a frequent 4 base cutter, to produce a set of random overlapping sequences between 15-23 kb, as described in Maniatis *et al.* (1982). The digested DNA was fractionated on a 0.4% agarose gel in 1 x TAE, and DNA between 15-23 kb was recovered by electroelution using the BIO-TRAP apparatus (Schleicher and Schule). The *Sau*3A ends of the genomic inserts were partially filled in with dATP and dGTP and cloned into the *Xho*I half-site arms of lambda GEM-11, a λ cloning vector capable of accepting 9-23 kb inserts, according to the supplier's instructions (Promega). Based on trial ligations, a ratio of 1:2, vector:insert DNA was used in the ligation reactions. λ DNA was packaged *in vitro* (Rosenberg, 1987) using *E. coli* KW 251 plating cells. Two genomic clones were recovered by screening with the 947 PCR product; the genomic inserts were cloned into a plasmid vector, pBS m13+.

The genomic inserts in plasmids p2A and p66 were sequenced from both ends initially (200 nucleotide). To sequence further into the insert, the nested deletion strategy was used to obtain a set of nested deletion subclones containing progressive unidirectional deletions of the genomic insert, for both plasmids p2A and p66 (Sambrook *et al.*, 1989). Several sets of nested deletion clones were generated from plasmids p2A and p66 using the Erase-A-Base system (Promega).

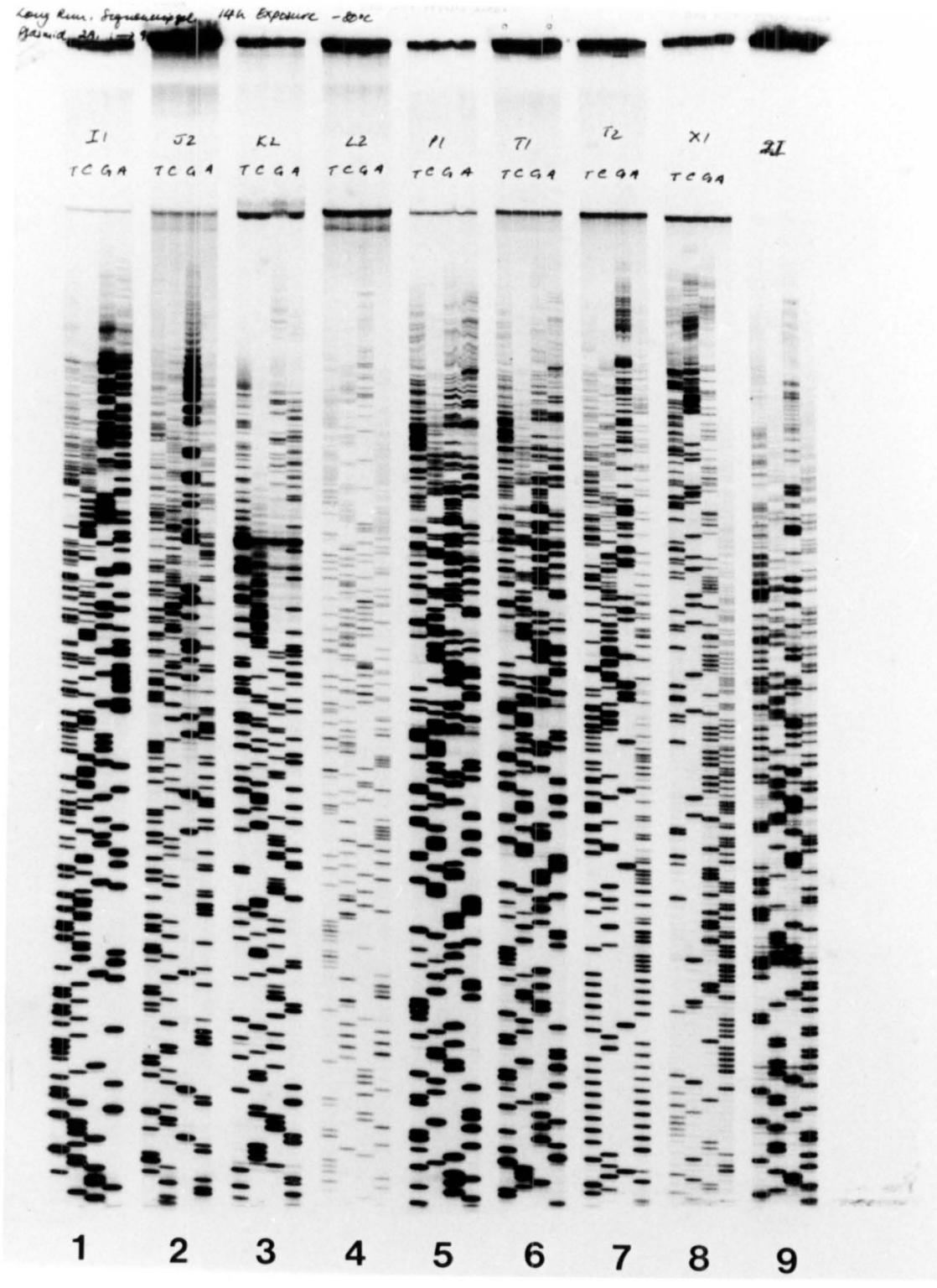
A2.2 Sequence analysis of plasmids p2A and p66

Gel electrophoresis of sequencing reactions tended to give more reliable results than automated sequencing of the templates. The results of automated sequencing was used to confirm stretches of overlap between 2 or more clones, rather than for obtaining new data. With manual sequencing, I was able to read up to 340 nt confidently. An example of a sequencing gel and a print out from the automated sequencer is given in Fig. A2.1 and A2.2

The sequenced regions within the genomic inserts in plasmids p2A and p66, are indicated in Fig. 3.2 (Chapter III) as 2A-1 to 2A-5, and 66-1 to 66-3, respectively. The relative position of the 947 PCR sequence in plasmids p2A and p66 was determined by DNA dot blot analysis (medium stringency) of the nested deletion clones generated for sequencing. Hybridization to the 947 PCR sequence was absent in plasmids less than 5.3 kb and 5.0 kb in size for p2A and p66 respectively, after 22 h exposure (data not shown), suggesting that the 947 PCR related sequence was found in the deleted regions of these plasmids, as indicated in Fig. 3.2 (Chapter III).

A search using DNASIS failed to locate the 947 PCR sequence within any of the sequenced regions of plasmids 2A and 66. Each sequence was individually searched for homology to the 947 PCR sequence, and also to peJK1 and peJK2 cDNA sequences, in both normal and complementary strand orientation, using DNASIS. In each search, both high, medium and low initial and optimized homology scores were examined for sequence identity. Examination for homology to the 947 PCR sequence, as well as the cDNA sequences, peJK1 and peJK2, in both strand orientations showed no regions > 25 nucleotides where the homologous nucleotides were contiguous. Generally, between 43-52% sequence identity was observed. The results of homology searches have been summarized (p2A, Tables A2.1 and 2; p66, Tables A2.3 and 4).

Figures A2.1. and A2.2 Examples of sequencing results obtained by manual sequencing and automated sequencing


Long Run. Sequencing gel. 14h Exposure -80°C
I1
J2
K1
L2
M1
T1
T2
X1
21
TCGA
TCGA
TCGA
TCGA
TCGA
TCGA
TCGA
TCGA
1
2
3
4
5
6
7
8
9

Model
Version 2.1.1

Jenny P66-R2 rev

Lane 17

Signal G:378 A:568 T:375 C:295
DyeTerminator{AnyPrimer}
591 MATRIX FILE
Points 644 to 8644 Base 1: 644

Fri, 24 Nov 1995 6:31 AM
Thu, 23 Nov 1995 3:44 PM
Spacing: 8.43 SemiAdaptive

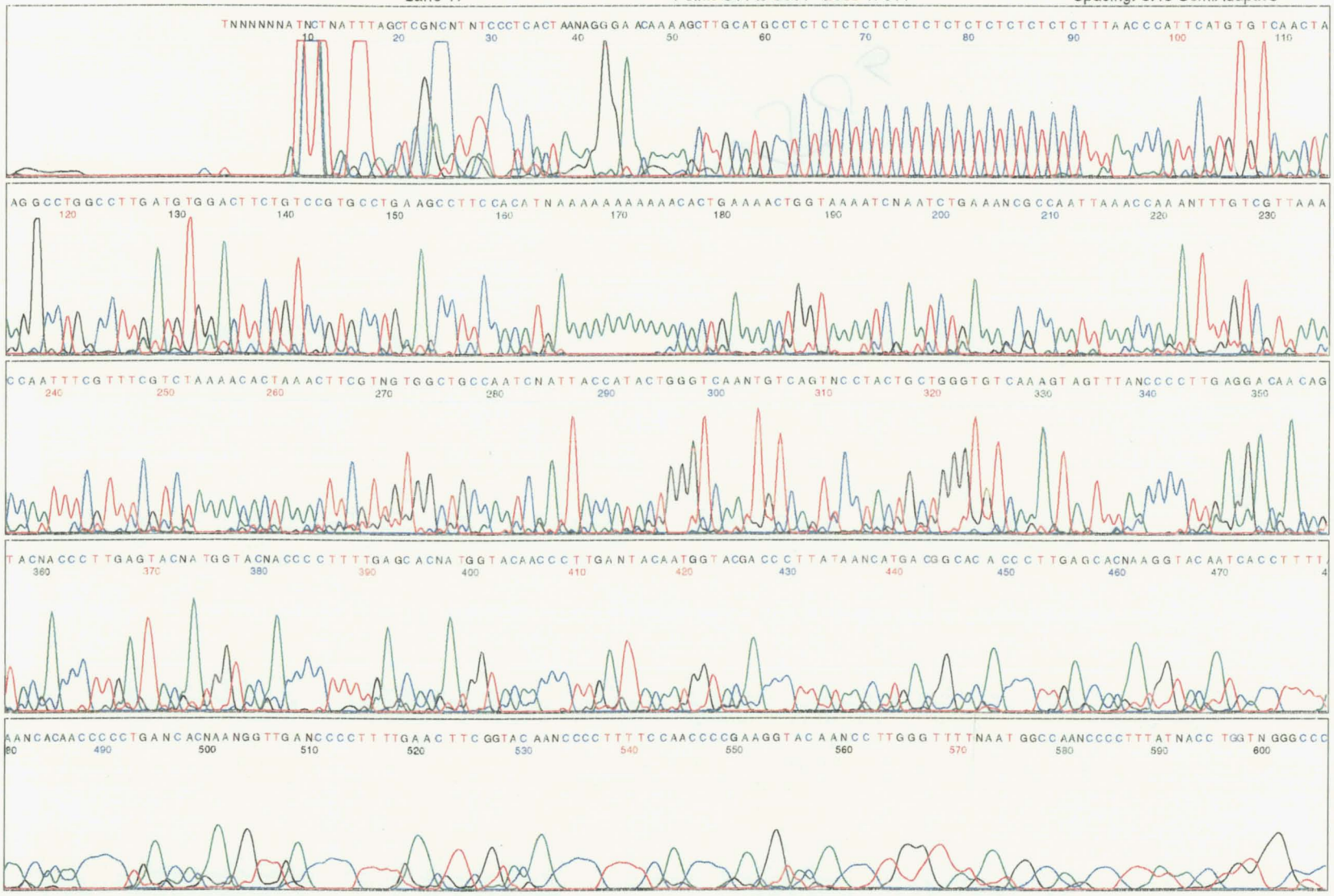


Table A2.1. Sequence identity (%) of regions within plasmid 2A, to the 947 PCR sequence in both normal and complementary (compl.) strand orientation. The similarity has been expressed in the % identity over the length of overlap between the two sequences.

Regions in Plasmid 2A	length of sequence	947 PCR (normal) % identity (bp overlap)	947 PCR (compl.) % identity (bp overlap)
2A-1	516	43.3% (446 bp)	46.2% (429 bp)
2A-2	722	44.0% (629 bp)	46.7% (606 bp)
2A-3	684	46.2% (582 bp)	44.5% (636 bp)
2A-4	1046	49.8% (257 bp)*	44.6% (766 bp)
2A-5	785	46.2% (688 bp)	49.5% (610 bp)

* Area within this sequence has ca. 25 homologous nucleotides which are contiguous.

Table A2.2. Sequence identity of regions within plasmid 2A, to cDNA sequences of peJK1 and peJK2, in both normal (N) and complementary (C.) strand orientation. The similarity has been expressed in the % identity over the length of overlap between the two sequences

Regions in Plasmid 2A	length of sequence	% identity (bp overlap) PeJK1 N.	peJK1 C.	% identity (bp overlap) PeJK2 N.	peJK2 C.
2A-1	516	47.1% (363)	48.8% (328)	46.5% (314)	45.7% (339)
2A-2	722	44.9% (543)	45.6% (579)	52.1% (238)	46.3% (350)
2A-3	684	46.5% (582)	45.2% (600)	42.8% (526)	43.3% (411)
2A-4	1046	46.7% (409)	49.5% (297)	47.4% (411)	45.4% (443)
2A-5	785	45.2% (504)	45.4% (471)	51.5 (225)	46.5% (538)

Table A2.3. Sequence identity (%) of regions within plasmid 66, to the 947 PCR sequence in both normal and complementary (compl.) strand orientation. The similarity has been expressed in the % identity over the length of overlap between the two sequences.

Regions in Plasmid 66	length of sequence	947 PCR (normal) % identity (bp overlap)	947 PCR (compl.) % identity (bp overlap)
66-1	1594	43.9% (935 bp)	52.5% (515 bp)
66-2	1878	44.7% (873 bp)	46.2% (739 bp)
66-3	894	42.9% (784 bp)	45.3% (830 bp)

Table A2.4. Sequence identity of regions within plasmid 66, to cDNA sequences of peJK1 and peJK2, in both normal (N) and complementary (C.) strand orientation. The similarity has been expressed in the % identity over the length of overlap between the two sequences.

Regions in Plasmid 66	length of sequence	% identity (bp overlap)		% identity (bp overlap)	
		PeJK1 N.	peJK1 C.	PeJK2 N.	peJK2 C.
66-1	1594	49.4% (577)*	47.8% (433)	44.6% (554)	56.5% (232)
66-2	1878	45.0% (5491	47.3% (347)	46.5% (525)	45.5% (552)
66-3	894	44.2% (389)	45.1% (517)	44.4% (576)	44.7% (450)

A2.3 Analyses of sequence data from plasmids p2A and p66

The sequenced regions of plasmids p2A and p66 both contained introns (non-coding sequences) as determined by examination of the open reading frames (ORFs) and protein coding region predictions based on Fickett's (1982) TESTCODE program. Possible transcription regulatory sequences such as the TATA box and polyadenylation sites have been highlighted in the sequence data of plasmids p2A (Fig. A2.3 to A2.7), and p66 (Fig. A2.8 to A2.10). However, in the absence of the definitive 947 PCR sequence and how the genomic clones relate to the cDNA clones, assignments of introns, exons and ORFs is speculative.

Sequence 2A-1

SEQUENCE : 516 BP; 133 A; 131 C; 107 G; 145 T.

```
           9          18          27          36          45          54
5' CGA TCG TAA CAC CCC CAA CAC TGG CAG GAG CCA CGA CTT TTC CTG CCG GAT GTT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Arg Ser *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    Asp Arg Asn Thr Pro Asn Thr Gly Arg Ser His Asp Phe Ser Cys Arg Met Phe
     Ile Val Thr Pro Pro Thr Leu Ala Gly Ala Thr Thr Phe Pro Ala Gly Cys Phe

          63          72          81          90          99         108
TTC CGG AGG AAG TGT TGG TTT GGA AGT TGC CTC CTG CTT TAC TAT CCC GTC TCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Ser Gly Gly Ser Val Gly Leu Glu Val Ala Ser Cys Phe Thr Ile Pro Ser Leu
  Pro Glu Glu Val Leu Val Trp Lys Leu Pro Pro Ala Leu Leu Ser Arg Leu Ser

         117         126         135         144         153         162
CTT CTC TAT AGA CTA GAA ACT TAG GTT ACT ATT ACT ATC CTG CAA CAG ATA ACC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ...
 Phe Ser Ile Asp *** ... ... ... ... ... ... ... ... ... ... ... *** ...
  Ser Leu *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

         171         180         189         198         207         216
TCG TTA TGG ACC ACC AGG TTC CCT GTT ATC TGT CTC TCC CAT GTA CCA GAC TTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... Met Tyr Gln Thr ***
  ... Met Asp His Gln Val Pro Cys Tyr Leu Ser Leu Pro Cys Thr Arg Leu Arg

         225         234         243         252         261         270
GGT TCA ACA CAC ACA ACC ATC TCC CTC AGC ATG AAC CTA ATT ATC TCT CTC AAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... Met Asn Leu Ile Ile Ser Leu Lys
 ... ... ... ... ... ... ... ... ... ... *** ... *** ... ... ... ... ...
  Phe Asn Thr His Asn His Leu Pro Gln His Glu Pro Asn Tyr Leu Ser Gln Val

         279         288         297         306         315         324
TTA ATT TCC GAG ATG GTC TGC CAT ACC TCA AGA TTA AGG CTT TTA TTA ATG TTG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Ile Ser Glu Met Val Cys His Thr Ser Arg Leu Arg Leu Leu Leu Met Leu
 *** ... ... ... ... ... ... ... ... ... ... *** ... ... ... *** ... ...
  Asn Phe Arg Asp Gly Leu Pro Tyr Leu Lys Ile Lys Ala Phe Ile Asn Val Gly

         333         342         351         360         369         378
GCG AAA AGA GCG CTG GCG GAA ATG GTT GCT GAA AAG GCA CTG ACA GTA ATA GTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Lys Arg Ala Leu Ala Glu Met Val Ala Glu Lys Ala Leu Thr Val Ile Val
 ... ... ... ... ... ... ... ... ... ... ... ... *** ... *** *** *** ***
  Glu Lys Ser Ala Gly Gly Asn Gly Cys *** ... ... ... ... ... ... ... ...

         387         396         405         414         423         432
GGT TGA GGC TTT ATT GGT TAC ACT CCT GGT TGC AAC AAT TGC CAG AAG GCT GGC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Gly *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

         441         450         459         468         477         486
GGG AAG GCG CTG CAC ACA CCC AGT TGG CGG CAC TCT GAA GAA AAC CAA TAT GTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Met Phe
  ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ...

         495         504         513
TTC AAT TTC AAA GCT ACT ACT TTG CCT CAA 3'
--- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ...
 Ser Ile Ser Lys Leu Leu Leu Cys Leu
  ... ... ... ... ... ... ... ... ...
```

Figure A2.3. Sequence data of genomic clone p2A, region 2A-1 (516 bp). The sequence data was analysed for open reading frame (ORF)s in three positions.

Sequence of 2A-2

SEQUENCE : 722 BP; 164 A; 212 C; 148 G; 198 T.

```
            9          18          27          36          45          54
5' AGC TCT CTA ATG AGT ACA GAA CGC ATT TCC TTC CAA TAT TTC ATC CCT GCA CTA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Ser Ser Leu Met Ser Thr Glu Arg Ile Ser Phe Gln Tyr Phe Ile Pro Ala Leu
    Ala Leu *** *** ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Leu Ser Asn Glu Tyr Arg Thr His Phe Leu Pro Ile Phe His Pro Cys Thr Thr

           63          72          81          90          99         108
   CCA AGC TCT CAC ACA CCA AAA TCT TCC CCA ACA TCT TAA GTC CTC AAG CTC TTC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Pro Ser Ser His Thr Pro Lys Ser Ser Pro Thr Ser *** ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Lys Leu Ser His Thr Lys Ile Phe Pro Asn Ile Leu Ser Pro Gln Ala Leu Pro

          117         126         135         144         153         162
   CAA ACA TTC TTA CGA TCT TAC TTC TAA ATA ACT CTA ATC ACT CTG ATA TTC AAA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... *** ... *** ... ... *** ... ... ...
     Asn Ile Leu Thr Ile Leu Leu Leu Asn Asn Ser Asn His Ser Asp Ile Gln Ser

          171         180         189         198         207         216
   GCA TCA TTC ATA TTT TCA GAA AAC GAA ACT TTT ACT TTG CTA CAC GTC TGA AGT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Ile Ile His Ile Phe Arg Lys Arg Asn Phe Tyr Phe Ala Thr Arg Leu Lys Ser

          225         234         243         252         261         270
   CTT CCA CTC TGC CTC CCA GGC TCT GGC TCA CAC TCT TAC TCA CAA AAC TAA AGC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Ser Thr Leu Pro Pro Arg Leu Trp Leu Thr Leu Leu Leu Thr Lys Leu Lys Pro

          279         288         297         306         315         324
   CAG ACT CCC CTG CTC TCT CCT ACT GAC ACA TGA CCT TTC AAA AGC TGC GGC CAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Asp Ser Pro Ala Leu Ser Tyr *** ... Met Thr Phe Gln Lys Leu Arg Pro Leu

          333         342         351         360         369         378
   TTC ATA CAT CAC GTC GCC AGC AAC ACA CTT CCC CTC TTA CCT CTT TCT CTA GTA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...
     His Thr Ser Arg Arg Gln Gln His Thr Ser Pro Leu Thr Ser Phe Ser Ser Thr

          387         396         405         414         423         432
   CAC CGC TCT ATT AGG GCA AGA TAT ATT AAC AGA CTG GGT CCA GGT TGT CTG TTA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ***
     Pro Leu Tyr *** ... ... ... ... *** ... ... ... ... ... ... ... ... ...

          441         450         459         468         477         486
   GTA GAC GGA TCG ACT AGG CCC AGC TCC TGT GGT AGC GGA GTG ATT GAC TAG GCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...
    *** ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ...
     ... ... ... ... *** ... ... ... ... ... *** ... ... ... *** ... ... ...

          495         504         513         522         531         540
   CAG CTC CTG TGG TAG CGG ACT GAT TGT CTG GGC CCA GCT CCT GTG GTA GCG GAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...
     ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

          549         558         567         576         585         594
   TGA TTG TCT GGG CCC AGC TCC TGT GGT ATC GGA CTG ATT GAC TGG GCC CAG CTC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ...
```

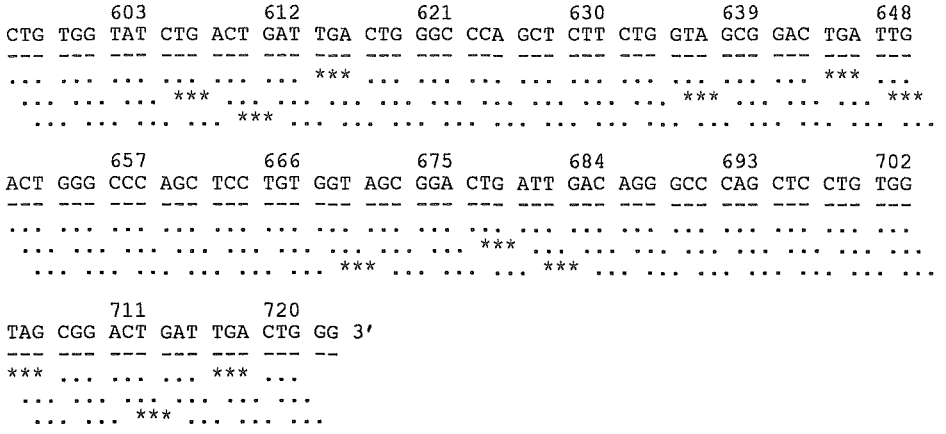


Figure A2.4. Sequence data from the genomic clone p2A, region 2A-2, (722 bp) was analyzed for open reading frames (ORF)s in all three positions. Several termination codons were located within this sequence (***).

Sequence 2A-3

SEQUENCE : 684 BP; 160 A; 142 C; 174 G; 208 T.

```
            9          18          27          36          45          54
5' AGG TTC CGT GCT TGA TTG AGT TTC CTT CCC CGG AAT GTC CAT CTT ACC TAG GAG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Arg Phe Arg Ala *** ... ... ... ... ... ... ... ... ... ... ... *** ...
    Gly Ser Val Leu Asp *** ... ... ... ... ... Met Ser Ile Leu Pro Arg Ser
     Val Pro Cys Leu Ile Glu Phe Pro Ser Pro Glu Cys Pro Ser Tyr Leu Gly Ala

            63          72          81          90          99         108
   CAA CCA GGT GCT CGG GGA TAC AAC AGG CGG GTC GCT CGG GGT CTA GGG TAG GGT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...
    Asn Gln Val Leu Gly Asp Thr Thr Gly Gly Ser Leu Gly Val *** ... ... ...
     Thr Arg Cys Ser Gly Ile Gln Gln Ala Gly Arg Ser Gly Ser Arg Val Gly Trp

           117         126         135         144         153         162
   GGA AGA TGT CAG GCT GTG AGC GGT TCT GGT TGG CGG GTA GAT TTT AGC CTG TGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ***
    ... ... ... ... ... *** ... ... ... ... ... ... *** ... ... ... ... ...
     Lys Met Ser Gly Cys Glu Arg Phe Trp Leu Ala Gly Arg Phe *** ... ... ...

           171         180         189         198         207         216
   AGT GCA GGA GTG GGT GGC ACC GCC CCA TTA CCT GTC CTC CTG CCC TTA CAG ACT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

           225         234         243         252         261         270
   TAG TGT GCT GTA GCA AAC GAG GTC TGA TGT TAA AGC ATG CTC TAC ATA CTG TAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   *** ... ... ... ... ... ... ... *** ... *** ... Met Leu Tyr Ile Leu Tyr
    ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... Met Leu Lys His Ala Leu His Thr Val His

           279         288         297         306         315         324
   ACT GTT GTG TTA TGT TCG TAT GGT AAG AAC ACA TGG TAG TAT GTA CGT CAC ATG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Thr Val Val Leu Cys Ser Tyr Gly Lys Asn Thr Trp *** ... ... ... ... Met
    ... ... ... ... ... ... Met Val Arg Thr His Gly Ser Met Tyr Val Thr Trp
     Cys Cys Val Met Phe Val Trp *** ... ... Met Val Val Cys Thr Ser His Gly

           333         342         351         360         369         378
   GTG TTT TGA ACA GTG TAC ACT CAG AAG TTG TCT CAC GCT ACT ACA GGT CTT ATT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Val Phe *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    Cys Phe Glu Gln Cys Thr Leu Arg Ser Cys Leu Thr Leu Leu Gln Val Leu Leu
     Val Leu Asn Ser Val His Ser Glu Val Val Ser Arg Tyr Tyr Arg Ser Tyr ***

           387         396         405         414         423         432
   AAA GTG TTC GCC TAA TAA GCT GAC TCT TTT TTT TCA CCT ACA CTA CCT ATA CAA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... *** *** ... ... ... ... ... ... ... ... ... ... ... ...
    Lys Cys Ser Pro Asn Lys Leu Thr Leu Phe Phe His Leu His Tyr Leu Tyr Asn
     ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

           441         450         459         468         477         486
   CAG GTA ACT ACC TTC ACC TTG GTG TAC CTA GCT GAT ATC AGG TTT AAC ACT ATG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Met
    Arg *** ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... *** ... ... ... *** ... ... ...

           495         504         513         522         531         540
   TCT ACC AGG GCC GGG GAC TGA CGT TGT CGA TAT ACT TGT AGG TAC GAG GAA TGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Ser Thr Arg Ala Gly Asp *** ... ... ... ... ... ... ... ... ... ... ***
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... Met Asn

           549         558         567         576         585         594
   ACT TAT CAT CCA CGG TAA TTG GGT GTC ATT TGG GCA GCC AAT GAT GAA GTT CAT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... Met Met Lys Phe Ile
     Leu Ser Ser Thr Val Ile Gly Cys His Leu Gly Ser Gln *** *** ... ... ...
```

```
        603         612         621         630         639         648
TTC ACT GCT TCT AAT GAC GTT CAA TAT CAG CCA GCC AAT GAG AGC ATT TGT TTC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Ser Leu Leu Leu Met Thr Phe Asn Ile Ser Gln Pro Met Arg Ala Phe Val Ser
  ... ... ... *** *** ... ... ... ... ... ... ... *** ... ... ... ... ...

        657         666         675         684
AGA GTC TTT GGA ATT ATA TGT AAT CTG TGT TGG TAA 3'
--- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ***
 Glu Ser Leu Glu Leu Tyr Val Ile Cys Val Gly
  ... ... ... ... ... Met *** ... ... ... ...
```

Figure A2.5. Sequence data from genomic clone p2A, region 2A-3 (684 bp), analysed for open reading frames in three positions. The location of the termination codons are indicated (***).

Sequence 2A-4

SEQUENCE : 1046 BP; 277 A; 305 C; 225 G; 239 T.

```
            9          18          27          36          45          54
5' AGG GCA CTC TGT ACC GGC TGA TGG TGG CAC AGA CGG CAG GGC CCT CCG TCC CGG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Arg Ala Leu Cys Thr Gly *** ... ... ... ... ... ... ... ... ... ... ...
    Gly His Ser Val Pro Ala Asp Gly Gly Thr Asp Gly Arg Ala Leu Arg Pro Gly
     Gly Thr Leu Tyr Arg Leu Met Val Ala Gln Thr Ala Gly Pro Ser Val Pro Ala

           63          72          81          90          99         108
   CCA ATG GTG GCA CAG ACG GCA GGG CCC TCC GTC CCG GCC GAT GGT GGC ACA GAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... Met Val Ala Gln Thr Ala Gly Pro Ser Val Pro Ala Asp Gly Gly Thr Asp
    Gln Trp Trp His Arg Arg Gln Gly Pro Pro Ser Arg Pro Met Val Ala Gln Thr
     Asn Gly Gly Thr Asp Gly Arg Ala Leu Arg Pro Gly Arg Trp Trp His Arg Arg

          117         126         135         144         153         162
   GGC AGG GCC CCC CGT ACC GGC AGA TGG TGG CAG AGA CGG CAG GAC CCT CCG TCA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Gly Arg Ala Pro Arg Thr Gly Arg Trp Trp Gln Arg Arg Gln Asp Pro Pro Ser
    Ala Gly Pro Pro Val Pro Ala Asp Gly Gly Arg Asp Gly Arg Thr Leu Arg His
     Gln Gly Pro Pro Tyr Arg Gln Met Val Ala Glu Thr Ala Gly Pro Ser Val Thr

          171         180         189         198         207         216
   CGT CCG CCA CTG TCT TGG AAA GGA GTG GAT GGA TTT TGA CCA AAC CTT TTG GGG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Arg Pro Pro Leu Ser Trp Lys Gly Val Asp Gly Phe *** ... ... ... ... ...
    Val Arg His Cys Leu Gly Lys Glu Trp Met Asp Phe Asp Gln Thr Phe Trp Gly
     Ser Ala Thr Val Leu Glu Arg Ser Gly Trp Ile Leu Thr Lys Pro Phe Gly Asp

          225         234         243         252         261         270
   ACT TGC TTT AAC CCC TGG AGC ACG ACG GTA TAA CCC TCA AGT ACG ACG GTA CGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ...
    Leu Ala Leu Thr Pro Gly Ala Arg Arg Tyr Asn Pro Gln Val Arg Arg Tyr Asp
     Leu Leu *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

          279         288         297         306         315         324
   CCC TTG AGC ACG ACA GTA CAG CCC TTA AGT ATC GTG GCC GGG GCT ATG ACC TGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... Met Thr ***
    Pro *** ... ... ... ... ... ... *** ... ... ... ... ... ... *** ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

          333         342         351         360         369         378
   CCC TGA GGG GTC AGT CAA GGA TCA TAT CAT ATA ATG CCC AAG GGC CAT ACC GTC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... *** ... ... ... ... ... ... ... ... ... Met Pro Lys Gly His Thr Val
    ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

          387         396         405         414         423         432
   GTG CTG AAG GGT CGA ACC GTC GTG CTC CAG GGG TTA AGG AGG GTC CAG TCA CCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Val Leu Lys Gly Arg Thr Val Val Leu Gln Gly Leu Arg Arg Val Gln Ser Pro
    ... *** ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

          441         450         459         468         477         486
   TAC TGG GTC AGA CTG CAT GGC TCT GAG GCA CTT GTA TCA AGT TAT GAG ACA CCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Tyr Trp Val Arg Leu His Gly Ser Glu Ala Leu Val Ser Ser Tyr Glu Thr Pro
    ... ... ... ... ... Met Ala Leu Arg His Leu Tyr Gln Val Met Arg His Pro
     ... ... ... ... ... ... ... *** ... ... ... ... ... ... *** ... ... ...

          495         504         513         522         531         540
   TCA CCC CTC CCC CAC CTC TCC CCT CCC CCA CCT CTC CCC TCC CCC ACC TCT CCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Ser Pro Leu Pro His Leu Ser Pro Pro Pro Pro Leu Pro Ser Pro Thr Ser Pro
    His Pro Ser Pro Thr Ser Pro Leu Pro His Leu Ser Pro Pro Pro Pro Leu Pro
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

          549         558         567         576         585         594
   CTC CCT CAC CCT TCC CCT CCC CCC CTC ACC ACT CCC CCA CGT CTC ACA CCC GAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Leu Pro His Pro Ser Pro Pro Pro Leu Thr Thr Pro Pro Arg Leu Thr Pro Asp
    Ser Leu Thr Leu Pro Leu Pro Pro Ser Pro Leu Pro His Val Ser His Pro Thr
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

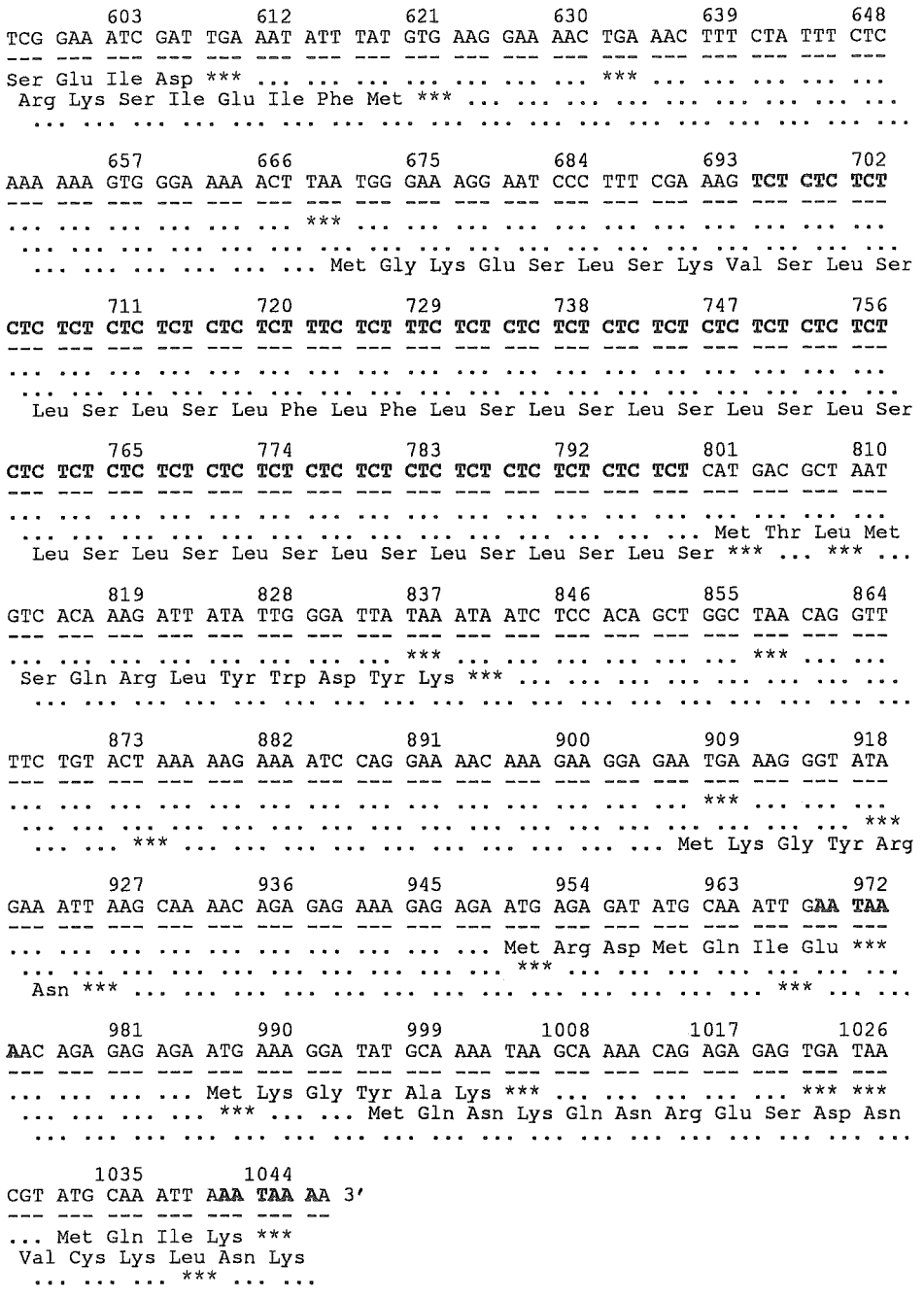


Figure A2.6. Sequence data from genomic clone p2A, region 2A-4 (1046 bp), analysed for open reading frames in three position. The 105 bp CT rich region (shown in bold), extending from nucleotides 694 to 798, has 98% sequence identity (100 bp overlap) to the CT rich region found in the C. **maenas** CHH cDNA precursor sequence. Possible polyadenylation sites are indicated (AATAAA). The location of the termination codons are indicated (***).

Sequence 2A-5

SEQUENCE : 785 BP

```
              9          18          27          36          45          54
5' ATG TAT GTA TGT ATG TAT GTA TGT ATA TAT GTA TGT ATG CGT ATG TAT GTA TGT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Met Tyr Val Cys Met Tyr Val Cys Ile Tyr Val Cys Met Arg Met Tyr Val Cys
    Cys Met Tyr Val Cys Met Tyr Val Tyr Met Tyr Val Cys Val Cys Met Tyr Val
     Val Cys Met Tyr Val Cys Met Tyr Ile Cys Met Tyr Ala Tyr Val Cys Met Tyr

             63          72          81          90          99         108
   ATG TAT GTA TGT ATA TAC GTA TAT TTG CAT TAG GTT GAG GGT CAC ACA CGT CCA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Met Tyr Val Cys Ile Tyr Val Tyr Leu His *** ... ... ... ... ... ... ...
    Cys Met Tyr Val Tyr Thr Tyr Ile Cys Ile Arg Leu Arg Val Thr His Val Gln
     Val Cys Met Tyr Ile Arg Ile Phe Ala Leu Gly *** ... ... ... ... ... ...

            117         126         135         144         153         162
   AAC ATT TTG TTT TTC CAG CCG TCA GTC ATC AGC AAA ACA TTT CGT CAG GTC TGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ***
    Thr Phe Cys Phe Ser Ser Arg Gln Ser Ser Ala Lys His Phe Val Arg Ser Glu
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

            171         180         189         198         207         216
   AAA CTC CTC ACA TCA ATT CTG AAG CTG ACT GAA GGG CAC TTA CCT CCG CTG CCA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    Asn Ser Ser His Gln Phe *** ... *** ... ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ...

            225         234         243         252         261         270
   GCT GCC GAA GTC ACG ACC CCT GCA ATA CCT GCC TGA ACC ACG GTA ATT AGA GAA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ...
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...

            279         288         297         306         315         324
   GAC ACA CTA CAC AAC TTC ATT AAT CTA CGC TCA TAT ATA TAT ATA TAT ACA CTG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

            333         342         351         360         369         378
   TAT ATA TAT AAA ATC AAA CAA TCA CTA CAA AAG AAT CCT AAT TCA ATT TAC TGA
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ***
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... *** ... ... ... ... ... ... ... ... ... *** ... ... ... ... ...

            387         396         405         414         423         432
   TTC ATA TAA ATA TGT ATC TCC TCC TTC AGT GTC GGC ACC TTC CCA GCC CCA TCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... Met Tyr Leu Leu Leu Gln Cys Arg His Leu Pro Ser Pro Ile Gln

            441         450         459         468         477         486
   AAC AAC CTG ACA GGC TAA CGC ACC TCA CTT CTA CCC TGC AGC CCC AGG GTT CCT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Gln Pro Asp Arg Leu Thr His Leu Thr Ser Thr Leu Gln Pro Gln Gly Ser Phe

            495         504         513         522         531         540
   TTG ATC TGT CGT CTA CTG CAG GTG CCT CGC CCT ACT TAT CGT TTT CAT CTT CTT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Asp Leu Ser Ser Thr Ala Gly Ala Ser Pro Tyr Leu Ser Phe Ser Ser Ser Ser

            549         558         567         576         585         594
   CAG TAC ACC TCT TCC TCC TCA GTC TAC CAT GTG CAT GTT CTC GTA CCT CCC AGC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    ... ... ... ... ... ... ... ... ... Met Cys Met Phe Ser Tyr Leu Pro Ala
     Val His Leu Phe Leu Leu Ser Leu Pro Cys Ala Cys Ser Arg Thr Ser Gln Pro
```

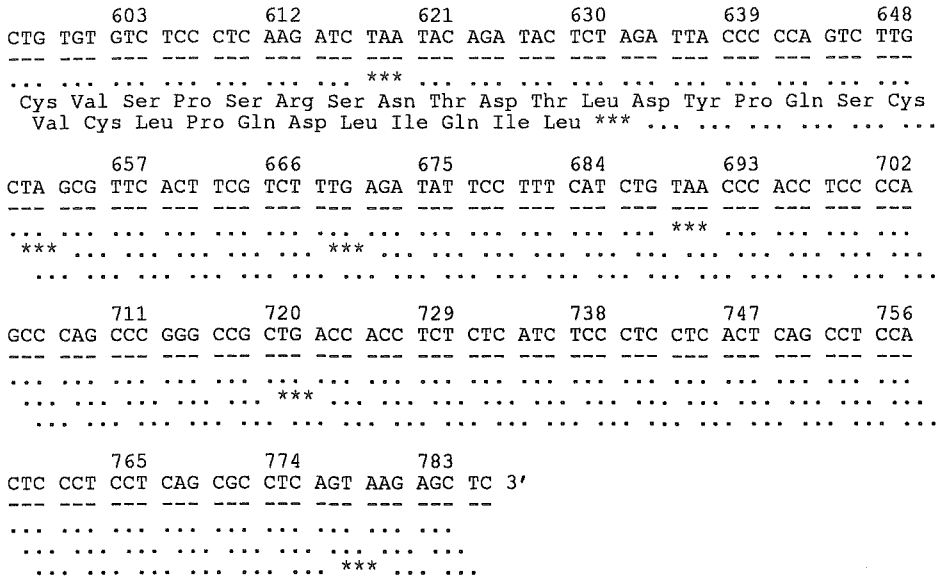


Figure A2.7. Sequence data from genomic clone p2A, region 2A-5 (785 bp), analysed for ORFs in three positions. A possible TATA box is highlighted. The termination codons are indicated (***).

Sequence 66-1

SEQUENCE : 1594 BP; 500 A; 416 C; 260 G; 418 T.

```
       549         558         567         576         585         594
TAA AAC GTT ACA CCG TCC TTA ACA CTT ATA ATC TAC TTA ACA AGA ACT TAT GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
*** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... *** ... ... *** ... ... *** ... ... ... Met Lys
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...

       603         612         621         630         639         648
ATT ATT TCT AGT ACA GAA AAC AAA TAC AAT GGA AAC GAT TTC CCT TCG TCA GCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Leu Phe Leu Val Gln Lys Thr Asn Thr Met Glu Thr Ile Ser Leu Arg Gln Gln
  ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

       657         666         675         684         693         702
GAT GCC TCG ACC ACC ATC AAA CAC ATA TCT TCA TTA CCT GCC CAC GTG TCC ATA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Met Pro Arg Pro Pro Ser Asn Thr Tyr Leu His Tyr Leu Pro Thr Cys Pro Tyr
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

       711         720         729         738         747         756
TCT ACA TTA AAC TGT ACA CCA GCT AAT TAT ATT CAT TAT AAT AAG CGA AAT AGC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Leu His *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... *** ... ... ... ... *** *** ... ... *** ...

       765         774         783         792         801         810
CAG CCA TTG TTC GGA AAT AGA GGT GTA GAT GAA TGG AAC AAA CTT GCA AGT GAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... *** Met Asn Gly Thr Asn Leu Gln Val Asn
  ... ... ... ... ... *** ... ... ... *** Met Glu Gln Thr Cys Lys *** ...

       819         828         837         846         855         864
CCA GTG AGT GCT GGA ACG GTT GGA ACT TGT AAG GGT AGG CTA GAT AAG GAT GGG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Gln *** ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... Met Gly
  ... ... ... ... ... ... ... ... ... *** ... *** ... ... *** ... ... ...

       873         882         891         900         909         918
CTA GAT AAG TAT ATG GAT AGG AAG GGA TGA CTT TAA CTT TAA GTC AGG GGC TGC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... Met Asp Arg Lys Gly *** ... *** ... *** ... ... ... ...
 *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... *** ... ... ... *** ... ... Met Thr Leu Thr Leu Ser Gln Gly Leu Pro

       927         936         945         954         963         972
CTA GCA TGG ACC AAT CGG CCT ACT GCA GTT TCT TTA TTT CTT ATG TTC GTA TGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... Met Phe Val Cys
 *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  Ser Met Asp Gln Ser Ala Tyr Cys Ser Phe Phe Ile Ser Tyr Val Arg Met Phe

       981         990         999        1008        1017        1026
TCT TAT GTT ATG TTA TGT TAT GTT CTT GTT CTT ATG TTC TTA TGT TTT TCT CAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ser Tyr Val Met Leu Cys Tyr Val Leu Val Leu Met Phe Leu Cys Phe Ser His
 ... Met Leu Cys Tyr Val Met Phe Leu Phe Leu Cys Ser Tyr Val Phe Leu Thr
  Leu Cys Tyr Val Met Leu Cys Ser Cys Ser Tyr Val Leu Met Phe Phe Ser Pro

      1035        1044        1053        1062        1071        1080
CTC TAC TAA CCA TCC TAC AAT ACA ATA GTT TAC ATA CAA GAG GCG TCA AAG TTT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Tyr *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Ser Thr Asn His Pro Thr Ile Gln *** ... ... ... ... ... ... ... ... ...
  Leu Leu Thr Ile Leu Gln Tyr Asn Ser Leu His Thr Arg Gly Val Lys Val ***

      1089        1098        1107        1116        1125        1134
GAC TTA GAA ATA TGA ATA ATA TAT TTT GAG AAA CTG ACA AAG CTC AGT GAA AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... *** ... ... ... *** ... ... ... ... *** ... ... ... ... ... ... ...
  ... ... ... Met Asn Asn Ile Phe *** ... ... ... ... ... ... *** ... ...

      1143        1152        1161        1170        1179        1188
TAT ATA AAT ACT ACA GTT GTC CCT TGC CGT TAA GAC TGC CTC GCG ACC TGA CGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... *** ...
 ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

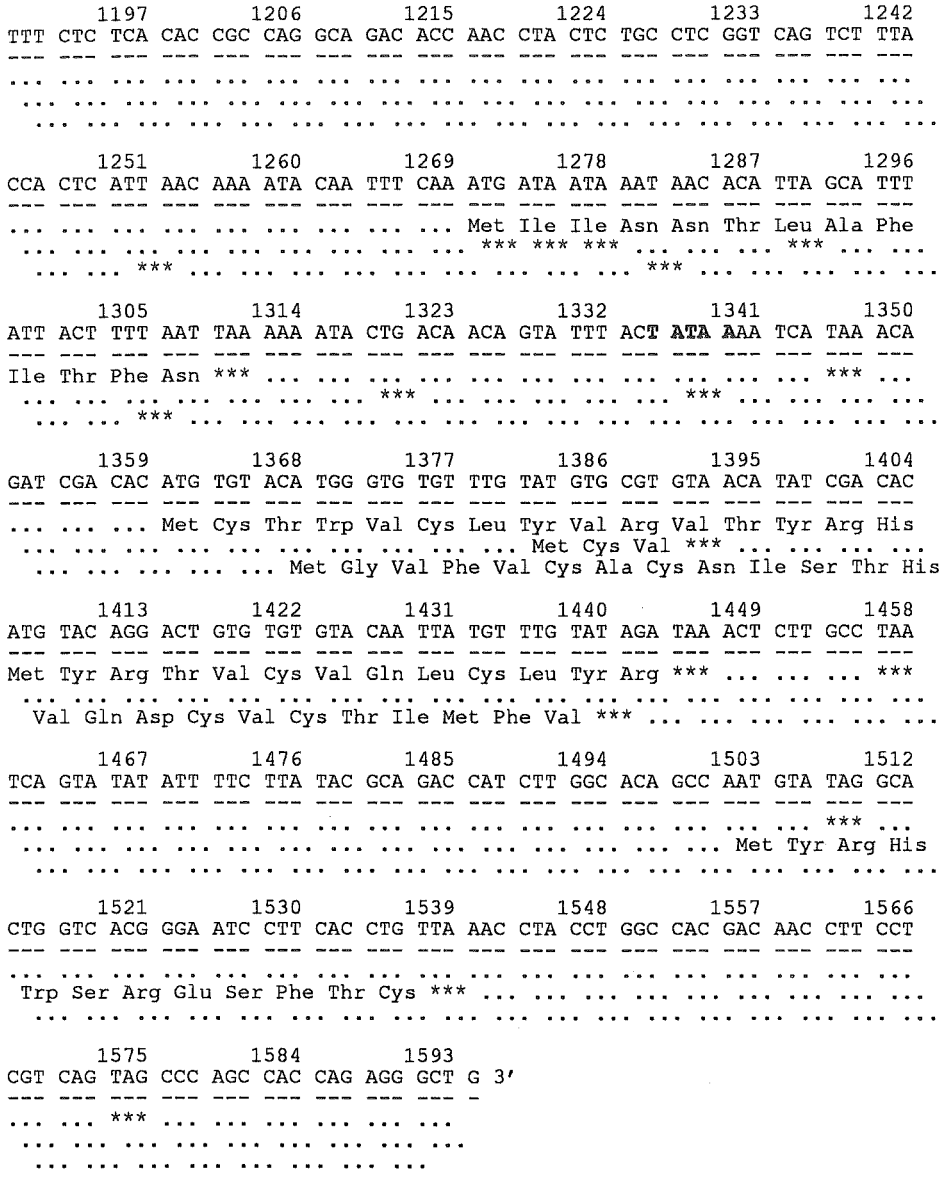


Figure A2.8. Sequence data of genomic clone p66, region 66-1, analysed for ORFs in three positions. A possible TATA box (highlighted) is located 25 bp upstream from the methionine initiation codon. The termination codons are indicated (***).

Clone 66-2

SEQUENCE : 1878 BP; 537 A; 469 C; 366 G; 506 T.

```
            9          18          27          36          45          54
5' TGC CCT CGT TAG ATG CGA ATT CGT TAG ATG TGA ACT CGT TAG ATG CGA GAT CTT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Cys Pro Arg *** Met Arg Ile Arg *** Met *** ... ... *** Met Arg Asp Leu
    Ala Leu Val Arg Cys Glu Phe Val Arg Cys Glu Leu Val Arg Cys Glu Ile Leu
     Pro Ser Leu Asp Ala Asn Ser Leu Asp Val Asn Ser Leu Asp Ala Arg Ser Tyr

           63          72          81          90          99         108
ACT GTA CCT TCC TTA CAT TGG ACG AAT ACA TCC CGG AAA GAT CTA TCC AAT CAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Thr Val Pro Ser Leu His Trp Thr Asn Thr Ser Arg Lys Asp Leu Ser Asn Gln
 Leu Tyr Leu Pro Tyr Ile Gly Arg Ile His Pro Gly Lys Ile Tyr Pro Ile Arg
  Cys Thr Phe Leu Thr Leu Asp Glu Tyr Ile Pro Glu Arg Ser Ile Gln Ser Gly

       117         126         135         144         153         162
GCA TAC CAA CTG ACC TAG CAC AAC CAA TTA CAG AAT ATG CCG CCC AGT TTC AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Tyr Gln Leu Thr *** ... ... ... ... ... ... Met Pro Pro Ser Phe Asn
 His Thr Asn *** ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  Ile Pro Thr Asp Leu Ala Gln Pro Ile Thr Glu Tyr Ala Ala Gln Phe Gln Tyr

       171         180         189         198         207         216
ATT CAT GGG ACA AAG TTC CCG AGG CCA CAC TGG CAT ACT GAT GAT GAG CAC CTC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ile His Gly Thr Lys Phe Pro Arg Pro His Trp His Thr Asp Asp Glu His Leu
 ... Met Gly Gln Ser Ser Arg Gly His Thr Gly Ile Leu Met Met Ser Thr Ser
  Ser Trp Asp Lys Val Pro Glu Ala Thr Leu Ala Tyr *** *** *** ... ... ...

       225         234         243         252         261         270
CCC GTA CTG TCC ATA ATG GAT GAC CTT CCA CTG ACA CTA ACA GGT CTC TTG CCC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Pro Val Leu Ser Ile Met Asp Asp Leu Pro Leu Thr Leu Thr Gly Leu Leu Pro
 Pro Tyr Cys Pro *** ... Met Thr Phe His *** ... *** ... ... ... ... ...
  ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

       279         288         297         306         315         324
TTG GCA TTA CTT TCC TTT GAC ATC ATC ATC TGG GGA GTC AAC GCT ACC TCT GTG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Ala Leu Leu Ser Phe Asp Ile Ile Ile Trp Gly Val Asn Ala Thr Ser Val
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ...

       333         342         351         360         369         378
TAC ACA CGG GGA AGA GTA CAC AAA CTT AAC AAT ACT TCT GTA CAT CAC CGG ACT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Tyr Thr Arg Gly Arg Val His Lys Leu Asn Asn Thr Ser Val His His Arg Thr
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ...

       387         396         405         414         423         432
ATT ATC CAG TTA CTT TTG TAC GTG TAC AGC ATA CAG CAT ACA GCA TAC AGC ATA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ile Ile Gln Leu Leu Leu Tyr Val Tyr Ser Ile Gln His Thr Ala Tyr Ser Ile
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

       441         450         459         468         477         486
CGG AAT ACA GCA TAC GGC ATC GAG GAA TGA CAT CAC AGC AGG TGG CGT AAT AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Arg Asn Thr Ala Tyr Gly Ile Glu Glu *** ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... Met Thr Ser Gln Gln Val Ala *** *** ...

       495         504         513         522         531         540
TCT CAG ATA TTA CAT ATC TAT TTA AGT TAG CCT ATG TCT TAC TGT GTC TGA CAG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... *** ... Met Ser Tyr Cys Val *** ...
 ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

       549         558         567         576         585         594
ACC ATG TGG CAT TTG TTT AAT AAA TAT AAA GTC TAC TCA CGC TGC TAT ACT ACG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... Met Trp His Leu Phe Asn Lys Tyr Lys Val Tyr Ser Arg Cys Tyr Thr Thr
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... *** *** ... *** ... ... ... ... ... ... ... ... ...
```

```
        603         612         621         630         639         648
CAC CAC CAT ACT AAT CCC CCC ATA GAG AAA CAA AAA TTT ACA TTA TTA CCG CCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
His His His Thr Asn Pro Pro Ile Glu Lys Gln Lys Phe Thr Leu Leu Pro Pro
 ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ...
  ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ...

        657         666         675         684         693         702
GCT TCT GCC GTT ACC CAG CTT GGC AAC TGC AAA AAT CCA ATA TTT ATT CCA CCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Ala Ser Ala Val Thr Gln Leu Gly Asn Cys Lys Asn Pro Ile Phe Ile Pro Pro
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

        711         720         729         738         747         756
GTG GGT CTG GGG GGT GGG GGT GAG GGA GGC AAG CAA TTT AGG TAA CAA AAT TTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Val Gly Leu Gly Gly Gly Gly Glu Gly Gly Lys Gln Phe Arg *** ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... *** ... ... ... ... ... *** ... ... ... ... ...

        765         774         783         792         801         810
CAT TCT TCT GCT TCT CTC GGC GTT ATA AAG CGT TAC CTA AAC AGA AAA CGT AAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... *** ... ... ... *** ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ...

        819         828         837         846         855         864
CGT CGT TAC CTA ACC TGT TAA AAG GAA GGC TCG GCC TCC CCC CTA ACC TCC CCG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... *** ... ... ... ... ... ... ... ... ... *** ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

        873         882         891         900         909         918
GAA CCC GCA GGC TAC ACA TAC AGT GAG AGT ATG TGT GGT TGT ATA GAT TTA ACT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... Met Cys Gly Cys Ile Asp Leu Thr
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... *** ...
  ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ...

        927         936         945         954         963         972
TGA ACA CCT GAA GAA GGT GTT CAT GTT ATT TCC ATT GGT GGT TAT ACC TTG ACG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
*** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... Met Leu Phe Pro Leu Val Val Ile Pro *** ...
  ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

        981         990         999        1008        1017        1026
TAT GCC TTA GGC CTT AAT GAC GAC CTA AGC TGC TGA TAG GCC AAA AAA CCA AAT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... *** *** ... ... ... ... ...
 Met Pro *** ... ... Met Thr Thr *** ... ... ... ... ... ... ... ... ...
  ... ... ... ... *** *** ... ... ... ... ... ... ... ... ... ... ... ***

       1035        1044        1053        1062        1071        1080
AAC CAA CCA ACC CGG AAC CCA GGC TAT GCT AAA TCT AAA ATT TAA TCT ACG TAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ***
 ... ... ... ... ... ... ... ... Met Leu Asn Leu Lys Phe Asn Leu Arg Asn
  ... ... ... ... ... ... ... ... ... ... *** ... *** ... ... ... ... ...

       1089        1098        1107        1116        1125        1134
TCA CCC GCG TAT ATG TCG CAC CTT TTT GCC CCC AAG TCA GGG CAA AAT TGG GCG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... Met Ser His Leu Phe Ala Pro Lys Ser Gly Gln Asn Trp Ala
 His Pro Arg Ile Cys Arg Thr Phe Leu Pro Pro Ser Gln Gly Lys Ile Gly Arg
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

       1143        1152        1161        1170        1179        1188
GTG CGG CTT ATA AAC GAG ATA GAC AGA GCG TAA TAT GCG GGT GAT TAC GCC GTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Val Arg Leu Ile Asn Glu Ile Asp Arg Ala *** ... ... ... ... ... ... ...
 Cys Gly Leu *** ... ... *** ... ... ... ... Met Arg Val Ile Thr Pro ***
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ...

       1197        1206        1215        1224        1233        1242
ATT AAG CAT GCG TCT CCG TAT TTC CAG ATG GTG ACC AAG ATT TAT CAT ATC AAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... Met Val Thr Lys Ile Tyr His Ile Asn
 ... ... Met Arg Leu Arg Ile Ser Arg Trp *** ... ... ... ... ... ... ...
  *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

```
      1251        1260        1269        1278        1287        1296
TAG AAA AGT GTC TAC GAA AAA AAT GCT GAA GTT TCA TGA GCA CAG CTG CTA TAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
*** ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ...
 ... ... ... ... ... ... ... Met Leu Lys Phe His Glu His Ser Cys Tyr Thr
  ... ... ... ... ... ... ... ... *** ... ... Met Ser Thr Ala Ala Ile Leu

      1305        1314        1323        1332        1341        1350
TTA GGC CAC AAC TGA CCT AAG CAA AAT AAC TTA ACC TCG TAC GGA CGG ACG GAC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ...
 *** ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ...
  Arg Pro Gln Leu Thr *** ... ... *** ... ... ... ... ... ... ... ... ...

      1359        1368        1377        1386        1395        1404
GGA CGG ACG GAC GTA CGG ACG GAG GGT AAC ATT TCA TGC CCC GCT CCA CTT GTA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... *** ... ... Met Pro Arg Ser Thr Cys Met

      1413        1422        1431        1440        1449        1458
TGG TTA CTC TTT ACT TCT GTC AGT GAA AAT TTA TTT CAC CTC TAC TGC AAT CAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  Val Thr Leu Tyr Phe Cys Gln *** ... ... ... ... ... ... ... ... ... ...

      1467        1476        1485        1494        1503        1512
ATA CAA CGA CGT CAA CTG TTT TTC TCC TCT CTC TCT CTC TCT CTC TCT CTC TCT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

      1521        1530        1539        1548        1557        1566
CTC TCT CTC TCT CTC TCT CTC TTT AAC CCA TTC ATG TGT CAA CTA AGG CCT GGC
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... Met Cys Gln Leu Arg Pro Gly
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ...
  ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ...

      1575        1584        1593        1602        1611        1620
CTT GAT GTG GAC TTC TGT CCG TGC CTG GAG CCT TCC ACA TGA AAA AAA AAA ACA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
Leu Asp Val Asp Phe Cys Pro Cys Leu Glu Pro Ser Thr *** ... ... ... ...
 ... Met Trp Thr Ser Val Arg Ala Trp Ser Leu Pro His Glu Lys Lys Lys His
  *** ... ... ... ... ... ... ... ... ... ... ... ... Met Lys Lys Lys Asn Thr

      1629        1638        1647        1656        1665        1674
CTG AAA ACT GGT AAA ATC GAA TCT GAA AGC GCC AAT TAA ACC AGA GTT TGT CGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ...
 *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  Glu Asn Trp *** ... ... ... *** ... ... ... ... ... ... ... ... ... ...

      1683        1692        1701        1710        1719        1728
TAA ACC AAT TTC GTT TCG TCT AAA ACA CTA AAC TTC GTT GTG GCT GCC AAT CGA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
*** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

      1737        1746        1755        1764        1773        1782
GTA CCA TAC TGG GTC AAG TGT CAG TTC CTA CTG CTG GGT GTC AGA GTA GTT TAA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ***
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

      1791        1800        1809        1818        1827        1836
CCC CTT GAG GAC GAC AGT ACG ACC CTT GAG TAC GAT GGT ACG ACC CTT TGA GCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...
 ... ... ... ... ... ... ... ... ... ... ... Met Val Arg Pro Phe Glu His
  ... *** ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ...
```

```
      1845        1854        1863        1872
CGA TGG TAC AAC CTT GAG TAC ATG GTA CGA CCT TAT AAG CAT 3'
--- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... Met Val Arg Pro Tyr Lys His
 Asp Gly Thr Thr Leu Ser Thr Trp Tyr Asp Leu Ile Ser
  Met Val Gln Pro *** ... ... ... ... ... ... *** ...
```

Figure A2.9. Sequence data from genomic clone p66, region 66-2 (1878 bp), analysed for ORFs in three positions. The CT rich region is indicated. A similar CT region was also found in the *C. maenas* CHH cDNA precursor. The termination codons are indicated (***).

Sequence 66-3

SEQUENCE : 894 BP; 203 A; 263 C; 179 G; 249 T.

```
            9          18          27          36          45          54
5' TTG CAT GCC TGC AGG TCG GTA CCG CCC TTT TCG AGC CCG ACG GTA CGA CCC TTG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Leu His Ala Cys Arg Ser Val Pro Pro Phe Ser Ser Pro Thr Val Arg Pro Leu
    Cys Met Pro Ala Gly Arg Tyr Arg Pro Phe Arg Ala Arg Arg Tyr Asp Pro Trp
     Ala Cys Leu Gln Val Gly Thr Ala Leu Phe Glu Pro Asp Gly Thr Thr Leu Gly

            63          72          81          90          99         108
   GGT ATG ATG GCA AGG CCT TTA TAA CCT GTG GGC TCC TAT CCC TAG CCT AAC ATG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Gly Met Met Ala Arg Pro Leu *** ... ... ... ... ... ... *** ... ... Met
    Val *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     Tyr Asp Gly Lys Ala Phe Ile Thr Cys Gly Leu Leu Ser Leu Ala *** ... ...

           117         126         135         144         153         162
   GCC AGG CCT TTA ACC TGT GGG CTC CTA TCC CTA GCC TAA CAT GGC CAG GCC TTT
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Ala Arg Pro Leu Thr Cys Gly Leu Leu Ser Leu Ala *** ... ... ... ... ...
    ... ... ... *** ... ... ... ... ... ... *** ... ... Met Ala Arg Pro Leu
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

           171         180         189         198         207         216
   ATA ACC TGT GGG CTC CTA TCC CTA GCC TAA CAT GGC CAG GCC TTT AAC CTG TGG
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ...
    *** ... ... ... ... ... ... *** ... ... Met Ala Arg Pro Leu Thr Cys Gly
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...

           225         234         243         252         261         270
   GCT CCT ATC CCT AGC CTA ACA TGG CCA GGC CTT TAT AAC CTG TGG GCT CCT ATC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
    Leu Leu Ser Leu Ala *** ... ... ... ... ... ... ... ... ... ... ... ...
     ... ... ... *** ... ... Met Ala Arg Pro Leu *** ... ... ... ... ... ...

           279         288         297         306         315         324
   CCT AGC CTA ACA TGG CCA GGC CTT TAA CCT GTG GGC TCC TAT CCC TAG CCT AAC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... *** ... ...
    ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
     *** ... ... Met Ala Arg Pro Leu Thr Cys Gly Leu Leu Ser Leu Ala *** ...

           333         342         351         360         369         378
   ATG GCC AGG CCT TTA ACC TGT GGG CTC CTA TCC CTA GCC TAA CAT GGC CAG GCC
   --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
   Met Ala Arg Pro Leu Thr Cys Gly Leu Leu Ser Leu Ala *** ... ... ... ...
    ... ... ... ... *** ... ... ... ... ... ... *** ... ... Met Ala Arg Pro
     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

```
        387         396         405         414         423         432
TTT ATA ACC TGT GGG CTC CTA TCC CTA GCC TAA CAT GGC CAG GCC TTT AAC CTG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ...
 Leu *** ... ... ... ... ... ... *** ... ... Met Ala Arg Pro Leu Thr Cys
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...

        441         450         459         468         477         486
TGG GCT CCT ATC CCT AGC TAA CAC CTT CCG TGT TCC CGT GGC CTG GCC TTT ACT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...
 Gly Leu Leu Ser Leu Ala Asn Thr Phe Arg Val Pro Val Ala Trp Pro Leu Leu
  ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ... ... ...

        495         504         513         522         531         540
ATC CTT GTA AAA ACC AAT CCC TTA CGG TTA TTC ATT CCC CTA TCC TTC TCT ATT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Ser Leu *** ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

        549         558         567         576         585         594
CAT TTA CAA ACA CAA AAA CGA GAA TGA CAT CAT TCA AAC CAT GCG TGT TCG TCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ...
 ... ... ... ... ... ... ... ... ... ... ... ... ... ... Met Arg Val Arg His
  ... ... ... ... ... ... ... Met Thr Ser Phe Lys Pro Cys Val Phe Val Thr

        603         612         621         630         639         648
CCG GAA GAA AGC TTA CGT CAC TGC ACA CCC ACT TGT ACA CTT TAC AAG ACA CGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Arg Lys Lys Ala Tyr Val Thr Ala His Pro Leu Val His Phe Thr Arg His Val
  Gly Arg Lys Leu Thr Ser Leu His Thr His Leu Tyr Thr Leu Gln Asp Thr Tyr

        657         666         675         684         693         702
ACA GCC AAG GAA TCA AGA ACC TTC CTC ACT TTG TCC CAT GAA TAT TGA AAC TCG
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ...
 Gln Pro Arg Asn Gln Glu Pro Ser Ser Leu Cys Pro Met Asn Ile Glu Thr Arg
  Ser Gln Gly Ile Lys Asn Leu Pro His Phe Val Pro *** ... ... ... ... ...

        711         720         729         738         747         756
GCG GAA TAT TCT GTA ATC GTT GAG CTG GTC AGT TGG TAT GCC TGA TTA GAT CGA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ...
 Arg Asn Ile Leu *** ... ... ... ... ... ... ... Met Pro Asp *** ... ...
  ... ... ... ... ... ... *** ... ... ... ... ... ... ... ... ... ... ...

        765         774         783         792         801         810
TCT TTC CCT GAT GTA TTC GTT CAA TGT GAG GAA GAG ACT TGA TTC CTT GAC TGT
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... *** ... ... ... ...
 ... ... ... Met Tyr Ser Phe Asn Val Arg Lys Arg Leu Asp Ser Leu Thr Val
  ... ... *** ... ... ... ... Met *** ... ... ... ... ... ... *** ... ...

        819         828         837         846         855         864
ACC TTC ACT GTC CAC AAT GTG GAT GAC TTT ACA AGA TGG TGT ACA AGT TGT TCA
--- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- --- ---
... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
 Pro Ser Leu Ser Thr Met Trp Met Thr Leu Gln Asp Gly Val Gln Val Val Gln
  ... ... ... ... ... ... ... *** ... ... ... Met Val Tyr Lys Leu Phe Arg

        873         882         891
GAG CCA TGA CGT GAA CAA CTG CTT GTT CAC 3'
--- --- --- --- --- --- --- --- --- ---
... ... *** ... ... ... ... ... ... ...
 Ser His Asp Val Asn Asn Cys Leu Phe
  Ala Met Thr *** ... ... ... ... ...
```

Figure A2.10. Sequence data from genomic clone p66, region 66-3 (894 bp), analysed for ORFs in three positions. Termination codons are indicated (***).

A2.4 Sequence comparison of cDNA clones with published sequences

The cDNA sequences were compared to published MIH and CHH cDNA sequences. Several examples are given below.

```
NO. TARGET FILE  DEFINITION                                                    INIT   OPT
 16 OCHHR.DNA                                                                   82   244
51.5% identity in 400 bp overlap

                    10         20         30         40         50         60
PEJK1.DNA      CACGAAGCGTGTGTAGTGCACAATCAGTTCCTACCACCACTACCACCTCTACCACCACCT
                                ::  :  ::::: : X:::::::::: : :: ::: :
OCHHR.DNA      ACCTGCCCTCTCCCACACGACACGCCCTTCCT-CAACCACTACCACCACCACTACCTGCC
              730        740        750        760        770        780
                    70         80         90        100        110
PEJK1.DNA      ACCACCTACCACCTACCACCACCACCACCACCACCTTCATCGTCATCTTCGC-TGTCGTC
                :   : ::::^v:::::::::::::: :::  :::  :   :: : :  : : :  : ::
OCHHR.DNA      TCTCTCCACCA-CTACCACCACCACCACAACCTGCTT-TTTTCCACCATGACTTACCTTC
               790         800        810        820        830        840
              120         130        140        150        160         170
PEJK1.DNA      -GCCT-CAAGACCTG-ATGCCCCCGGAGCCACAGCTCGTC-AGGA--AGGTCCTCCCAGC
                 ::: :: ::: ::  : : :::  : : ::  : :  : ::::   : ::::: :  :
OCHHR.DNA      TCCCTCCACGACTTGCCTTCTCCCTTAACGACGACACCACTAGGATCTGATCCTCTCCCC
                 850        860        870        880        890        900
                    180        190        200        210        220        230
PEJK1.DNA      A-CAGGGAAG-GTCTTCAAAGGCCTCATCCGTCATCATCACCACAGTCATTCACATGTTA
               : ::    ::  ::: : :    : :   :  :::::: :::::   :: : :::   ::
OCHHR.DNA      ACCAATACAGCCTCTACCATCTGCCCTATCTCCATCATTACCACCACCACT-ACAACCTA
                 910        920        930        940        950         960
                   240        250        260        270        280
PEJK1.DNA      AC-AAGTCAACAACA-AACCCAGATGGCCCATTCATGACACTTTCTTTGGTTCTGCAGCT
                :    ::  :: ::  ::::  :   :    :  : : : : :     : :  :  :
OCHHR.DNA      CCGTTCTCTCCACCACTACCCCCACCACGGCCTGCCGTCTC-TCCACCACTACCACCACC
                 970        980        990       1000       1010       1020
              290        300        310        320        330        340
PEJK1.DNA      TTGAAGATTTGTTTCTGTGATATTATTTCTCTCTTTTGTATATCCATTTTTGGTTTTCTC
                 :       :: :::     : ::     ::: :  : :  ::      :   ::  : :
OCHHR.DNA      ACGGCCTGCCGTCTCTCCACCACTAGGCACCTC-TGAGCACCTCATACATCTCTTGGCAC
                1030       1040       1050       1060       1070       1080
              350        360        370        380        390        400
PEJK1.DNA      ATACGTTGACTGGGTTGGTCT-GATTCAGAAGTGTAGTGAAGATAAG-GATATATATGAA
               : : :   : :   : ::  : :::  : :  : : : :  : : :  :  :::: :  ::
OCHHR.DNA      AGATGAACA-TATATGGGCATAGATGGACACATCTGGGCACGTTTTGCCATATCTGGGAC
                  1090       1100       1110       1120       1130       1140
               410        420        430        440        450        460
PEJK1.DNA      CTTGACTTTAGGATTCTATTGAGTTGGTTCTGTTATTCATAAATGTTCTTCTCACTCTAA

OCHHR.DNA      AGTTAGACACATTTGGGCACATTAACAAAGCGACTGGCACAGCCGATACAACTGGGTACA
                    1150       1160       1170       1180       1190       1200
               470        480        490        500        510        520
PEJK1.DNA      GTTCCTCTCGTCTGTTCTTGCCATCATTATCTCATGCTCACGAGGTATTTTTCTGTGGAC

OCHHR.DNA      ATTGACACATCTGGGCAAAGTTAGACTATGGCCGCCAACGGACCCCAAATGATACCAACG
                    1210       1220       1230       1240       1250       1260
               530        540        550        560        570        580
PEJK1.DNA      ATTAGAATGTATATTGGCCAAAATATATGAAAGAATTGTATAAAAAAAAAAAAAAAAAA

OCHHR.DNA      GACGCCGGTGGCATGCCTGCAGACGCTCGTAGATACCAACGGACGCCGGTGGATGCCT
                    1270       1280       1290       1300       1310
```

Figure A2.11. Sequence comparison of the peJK1 cDNA sequence with the crustacean hyperglycemic hormone A* (CHH-A*) from the crayfish , *O. limosus* (de Kleijn D.P.V. *et al.*, 1994a). peJK1 shared 51.5% identity in 400 bp overlap with the CHH cDNA sequence. The coding sequence for CHH-A* is from 106 to 513. In the figure, CHH-A* is indicated as OCHHR.DNA. The initial and optimized scores on DNASIS are indicated on the left.

```
NO. TARGET FILE  DEFINITION                                                    INIT   OPT
 48 SUNCMIH.SEQ                                                                 26   158
47.6% identity in 490 bp overlap

                     130       140       150       160       170       180
PEJK1.DNA      CCTCAAGACCTGATGCCCCCGGAGCCACAGCTCGTCAGGAAGGTCCTCCCAGCACAGGGA
               : : :  : ::: :: :  :   : :     :::::::::    :: ::::    :  :::
SUNCMIH.SEQ    CTTGAGAAGCTGCTGTCGTCCTCGTC-GTCTTCGTCAGGCTCTTCTTCCC---CCCTGGA
                     10        20        30        40            50
                      190       200       210       220       230
PEJK1.DNA      AGGTCT--TCAAAGGCCTCATCCGTCATCATCAC--CACAGTCATTCACATGTTAACAAG
               : :::   :   : :: :: :    :  : : :  :::  ::   ::: :  :   : :
SUNCMIH.SEQ    TGCTCTCGGCGGCGACCACAGCGTGAACAAGCGCGACACCTTCGACCAC-TCCTGCAAGG
                 60        70        80        90       100       110
                240       250       260       270       280       290
PEJK1.DNA      TCAACAACAAACCCAGATGGCCCATTCATGACACT-TTCTTTGGTTCTGCAGCTT-TGAA
               :: : ::  :::  ::  :: : ::::  :  ::   :   :  : ::  : ::   :
SUNCMIH.SEQ    GCATCTAC-GACCGGGA--GCTC-TTCAGAAAGCTGGACCGCGTCTGTGAGGATTGCTAC
                 120       130       140       150       160       170
                 300       310       320       330       340       350
PEJK1.DNA      GATTTGTTTC-TGTG-ATATTATTTCTCTCTTTTGTATATCC-ATTTTTGGTTTTCTCAT
               :  :::::: : :   :  :  :       :: :  ::: ::: :   :   : : :
SUNCMIH.SEQ    AACGTGTTTCGCGAGCCCAAGGTGGCCACGGAGTGCAAGTCCAATTGCTTCGTGAATAAG
                     180       190       200       210       220       230
                     360       370       380       390       400
PEJK1.DNA      ACGTTGACTG--GGTTGGTCTGAT-TCAGAAGTGTAGTGAAGATAAGGATATATATGAAC
               : ::: : ::   ::  : ::::: :::::  ::  :: :        :  : :::
SUNCMIH.SEQ    AGGTTCAATGTCTGTGTGGCTGATCTCAGACATGATGTCAGCCGCTTTCTGAAAATGGCT
                     240       250       260       270       280       290
               410       420       430            440       450       460
PEJK1.DNA      TTGACTTTAGGATTCTATTGAGT--TGGTTCT---GTTATTCATAAATGTTCTTCTCACT
                   ::   v :: ^:: ::  v   :: ^ :   :: : :  ::      :::: : :
SUNCMIH.SEQ    AATTCTGCGCTATCCTAATGGTTGAAGGCTATGGAGTAACTGCTACGCCAACTTCGTATT
                     300       310       320       330       340       350
                     470       480       490       500       510       520
PEJK1.DNA      CTA--AGTTCCTCTCGTCTGTTCTTGCCATCATTATCTCATGCTCACGAGGTATTTTTC-
               : :  ::: ::::     : : :::    ::   ::  :::  : :::::    :   :
SUNCMIH.SEQ    CAAGCAGTGCCTCGACGATCTCCTTATGGTC--GATGCCAT--TGACGAGTACGTGAACA
                     360       370       380       390       400
                       530       540       550       560       570
PEJK1.DNA      -TGTGGACATTAGAATGTA-TATTGGCCAAAATATAT---GAAAGAATTGTATAAAAAAA
                 ::  :  :   :  : : ::  :::  :: : : :   ::    : :::  :  :::
SUNCMIH.SEQ    CCGTCCAGCTGGTAGGGAAGTAAAGGCAGAAGTCTCTCAGGACGCTAATGTGGAGGAAAC
               410       420       430       440       450       460
                 580
PEJK1.DNA      AAAAAAAAAA
               :: :::::::
SUNCMIH.SEQ    AAGAAAAAAA
               470
```

Figure A2.12. Sequence comparison of peJK1 with the putative MIH cDNA sequence (SUNCMIH) from the shrimp, *P. vannamei* (Sun, 1994), using DNASIS. Both sequences share 47.6% identity over a 490 bp overlap region.

```
NO. TARGET FILE  DEFINITION                                                    INIT   OPT
  1 SUNCMIH.SEQ                                                                 68   168
46.1% identity in 436 bp overlap

                     100       110       120       130       140       150
PEJK2.DNA      GAGCCAGAGCGTGAGCCCGGCCGCAGATATGGTGGTGGCGGCTTCGGTGGTGGTGGTGGT
                                                                        : :: :
SUNCMIH.SEQ    CTTGAGAAGCTGCTGTCGTCCTCGTCGTCTTCGTCAGGCTCTTCTTCCCCCCTGGATGCT
                    10        20        30        40        50         60
                    160       170       180       190       200       210
PEJK2.DNA      GGTGGCGGCTTCGGTGGTGGCCACCATGGAGGCAGCTTCGGTGGTGGCAGCTTCGGTGGT
                  ::::::  :    : :   :: :  : : :: ::::      : :  :: : :
SUNCMIH.SEQ    CTCGGCGGCGACCACAGCGTGAACAAGCGCGACACCTTC------GACCACTCCTGCAAG
                    70        80        90             100       110
                   220       230       240       250       260
PEJK2.DNA      GGCAGTTTCG-GTGGTGGCCGTCGTGGTGGCTATGGCGGCTACCGTGGCTAGAGGTAACA
               :::: : ::   ::  ::  :   :   ::: : : ::  : :::   :  : : :::
SUNCMIH.SEQ    GGCATCTACGACCGGGAGCTCTTCAGAAAGCT-GGACCGCGTCTGTGAGGATTGCT-ACA
                  120       130       140       150       160       170
             270       280       290       300       310       320
PEJK2.DNA      AACTGGAT-GTGCG-GCAGGACGGACA-GACGAACAAGATC--TACCCACGAGAACCAGC
               :  ::  : : : :  :: :  :: :: :  :  ::::  :  :  :  :: :::  ::
SUNCMIH.SEQ    ACGTGTTTCGCGAGCCCAAGGTGGCCACGGAGTGCAAGTCCAATTGCTTCGTGAATAAGA
                   180       190       200       210       220       230
                  330       340       350       360       370       380
PEJK2.DNA      AAGACGACGAGCGTGTGGCAGTGACGTCACTCCTG--CACGACAACGTCCAG-GCGTCTC
                 : : :   :::::::  :::  :::  : ::    :  :  : : : :    :  :
SUNCMIH.SEQ    GGTTCAATGTCTGTGTGGC--TGATCTCAGACATGATGTCAGCCGCTTTCTGAAAATGGC
                  240       250       260       270       280       290
                   390       400       410       420       430       440
PEJK2.DNA      CAGTTTTCAACTCGTCCCTTAACTACAACCTTAAAGGAATCTCCACACATTAAGATGTAA
                : :: :   ::   :   :   :  :  ::     : : :: :  ::   ::  :   :
SUNCMIH.SEQ    TAATTCTGCGCTATCCTAATGGTTGAAGGCTATGGAGTAACTGC-TACGCCAACTTCGTA
                   300       310       320       330       340
                   450       460       470       480       490
PEJK2.DNA      TTCAAGC-GATCTTTTGC-ATTTGCTTAATGTTTCT-CCTTAAGTGAATATTTGTTTTTA
               ::::::: :  : :   : :: : ::::  ::   : :: :    :: ::  ::
SUNCMIH.SEQ    TTCAAGCAGTGCCTCGACGATCTCCTTATGGTCGATGCCATTGACGAGTACGTGAACACC
             350       360       370       380       390       400
              500       510       520       530       540       550
PEJK2.DNA      TAAGATTTGTAATATCTCAAAATCAAATAAACAAGAAACTGGAAGATTTCGAAAAAAAAA
                  :  ::  :      :::   :  :  :    :   :   :: : ^: v::: ::
SUNCMIH.SEQ    GTCCAGCTGGTAGGGAAGTAAAGGCAGAAGTCTCTCAGGACGCTAATGTGGAGGAAACAA
             410       420       430       440       450       460
              560
PEJK2.DNA      AAAAAAAAAA
                ::::::::^
SUNCMIH.SEQ    GAAAAAAAAA
             470
```

Figure A2.13. Sequence comparison of peJK2 cDNA sequence with the putative MIH cDNA sequence (SUNCMIH) from the shrimp, *P. vannamei* (Sun, 1994), using DNASIS. Both sequences share 46.1% identity over a 436 bp overlap region.

```
NO. TARGET FILE  DEFINITION                                                INIT   OPT
   1 CMCHH.DNA                                                              88   176
 45.4% identity in 509 bp overlap

                    10        20        30        40        50        60
 PEJK2.DNA      CGAAGAACACCAAGACCCCAGCCAACATGAACTTTCTGCGAGCTCTGCTGATGATGGCGG
                                                                X::  ::::::
 CMCHH.DNA      TCCCGCACGCACACGCACGCTCCACGCAAGGCTACGGACGCATGGATAGGATTCTGGCGG
                 140       150       160       170       180       190
                       70        80        90       100       110
 PEJK2.DNA      CCGTGATGGCATGGTGTACGTGGCCCTGGCGAGCCAGAGC-GT-GAGCCCGG--CCGCAG
                :: :::   : : :   : :  :::: :   ::::  :::v^: :::   ::  :  ::
 CMCHH.DNA      CCTTGAAAACCTCGCCAATGGAGCCCAGCGCAGCCCTAGCGGTGGAGAATGGAACTACAC
                 200       210       220       230       240       250
                 120       130       140       150       160       170
 PEJK2.DNA      ATATGGTGG--TGGCGGCTTCGGTGGTGGTGGTGGTGGCGGCTTCGGTGGTGGCCAC
                :   : :::    : :::     :        ::  ::   : :  ::       :  : :
 CMCHH.DNA      ACCCGTTGGAAAAGAGGCAAATTTACGACACGTCCTGCAAGGGTGTTTACGACCGTGCTC
                 260       270       280       290       300       310
                   180       190       200       210       220       230
 PEJK2.DNA      CATGGAGGCAGCTTCGGTGGTGGCAGCTTCGGTGGTGGCAGTTTCGGTGGTGGCCGTCGT
                  :  :   : ::: :     :   :   :: : ::  ::   ::    :   :     :
 CMCHH.DNA      TGTTCAATGA-CTTGGAGCACGTGTGTGACGATTGTTACAACCTCTACAGAACCTCCTAT
                 320       330       340       350       360       370
                   240       250       260       270       280       290
 PEJK2.DNA      GGTGGCTATGGCGGCTACCGTGGCTAGAGGTAACAAACTGGATGTGCGGCAGGACGGACA
                : :: ::  : : ::         :: :   :: :::  :::   : :::::  : : :
 CMCHH.DNA      GTTGCCTCGGCCTGCAGATCAAACT-GCTATAGCAACTTGGTGTTCCGGCA--ATGCATG
                  380       390       400       410       420         430
                   300       310       320       330       340       350
 PEJK2.DNA      GACGAAC-AAGATCTACCCACGAGAACCAGCAAGACGACGAGCGTGTGGCAGTGACGTCA
                :: :: :    :   :   :::::    : ::: : : : ::   :   :::     :
 CMCHH.DNA      GATGACCTTTTAATGATGGACGAGTTTGACCAATATG-CCAGAAAGGTACAGATGGTTGG
                     440       450       460       470       480       490
                  360       370       380       390       400       410
 PEJK2.DNA      CTCCTGCACGACAACGTCCAGGCGTCTCCAGTTTTC-AACTCGTCCCTTAAC--TACAAC
                :     : :  :::  :   :    : :     : ::: :   :: : ::  :::::
 CMCHH.DNA      CAGGAAGAAGTAAACAACAGATCTCAACAAACAAACAAACACCCGCCATCACTCTACAAG
                      500       510       520       530       540       550
                       420       430       440       450       460
 PEJK2.DNA      CTTAAAGGAATCTCCACACATTAAG-ATGTAATTCAAGCGATCT-TTTGCATTTGCTTAA
                :  ::    : :  ::::::   :: :  :: :: ::  : ::: : :  :::    ::
 CMCHH.DNA      C-CAATCACACCAACACACACACAGAACTTATTTTAAAAGGTCTCTATTTATTGATCTAG
                       560       570       580       590       600       610
               470             480       490       500       510       520
 PEJK2.DNA      TGTT------TCTCCTTAAGTGAAT--ATTTGTTTTTATAAGATTTGTAATATCTCAAAA
                : ::      ::::::   :::  :  :   :    :  :: : :::   : : : : :
 CMCHH.DNA      TCTTCCTAAGTCTCCTCGCGTGCCTCCAGGGGAGAGTTCAACAGTTG--CTTTATAATAT
                      620       630       640       650         660
                      530       540       550       560
 PEJK2.DNA      TCAAATAAACAAGAAACTGGAAGATTTCGAAAAAAAAAAAAAAAAAAAA
                  :     ::::::: ::
 CMCHH.DNA      GTACTCTCACAAGAATCTCTCACAACCTCTCACAATGGCTGGATATTA
               670       680       690       700       710
```

Figure A2.14. Sequence comparison of the peJK2 cDNA sequence with the CHH precursor cDNA sequence from the crab, *C. maenas* (Weidemann *et al.*, 1989). The coding sequence of this gene is from nucleotides 76 to 504; nucleotides 76 to 153 codes for the signal peptide, and 275 to 501, the mature peptide. peJK2 shared 45.4% identity in 509 bp overlap with the CHH cDNA sequence (CMCHH.DNA).